

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 22:31:05 ; Search time 3253 Seconds
(without alignments)
286.286 Million cell updates/sec

Title: US-09-895-435-3

Perfect score: 32

Sequence: 1 taatacacacaataattattgtgtgattatta 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_ov.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_mu.*
21: em_or.*
22: em_ov.*
23: em_ph.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rtd.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	32	100.0	4149	1	BTTN4430	X07651 Bacillus th
2	32	100.0	9672	1	BTPG12XX	X13481 Bacillus th
C 3	25.6	80.0	4149	1	BTTN4430	X07651 Bacillus th
C 4	25.6	80.0	9672	1	BTPG12XX	X13481 Bacillus th
C 5	24.6	76.9	164303	2	AC027541	AC027541 Homo sapi
6	24.6	76.9	176759	2	AC100807	AC100807 Homo sapi
7	23	71.9	81885	10	AL663040	AL663040 Mouse DNA
C 8	23	71.9	94535	9	AC083886	AC083886 Homo sapi
9	23	71.9	115359	8	AC003028	AC003028 Arabidops
C 10	23	71.9	119819	9	AL161631	AL161631 Human DNA
C 11	23	71.9	123579	2	AC017792	AC017792 Drosophi
C 12	23	71.9	145734	2	AC011096	AC011096 Homo sapi
C 13	23	71.9	151535	2	AC084188	AC084188 Homo sapi
C 14	23	71.9	164982	3	AC008254	AC008254 Drosophi
C 15	23	71.9	176036	3	AC009258	AC009258 Drosophi
C 16	23	71.9	194905	9	AC068896	AC068896 Homo sapi
C 17	22.6	70.6	102507	9	HSJ324N14	HSJ324N14 Human DNA
C 18	22.6	70.6	157564	2	AC025996	AC025996 Homo sapi
C 19	22.6	70.6	170073	2	AL845419	AL845419 Danio rer
C 20	22.6	70.6	171179	2	AC034140	AC034140 Homo sapi
C 21	22.6	70.6	182740	2	AC121330	AC121330 Homo sapi
C 22	22.6	70.6	198742	2	AL355360	AL355360 Homo sapi
C 23	22.4	70.0	363	8	MISCARS	X00436 Yeast mtDNA
24	22.4	70.0	676	8	YSCMTAP92	J01461 Yeast (S.ce
25	22.4	70.0	909	8	MISC28	X00818 Yeast mitoc
26	22.4	70.0	3542	8	YSCMTG16	L36900 Saccharomyc
27	22.4	70.0	33507	7	AY082070	AY082070 Bacteriop
28	22.4	70.0	85779	8	SCE011856	AC011856 Saccharom
29	22.4	70.0	91061	2	AC087139	AC087139 Mus muscu
C 30	22.4	70.0	132000	2	AC116976	AC116976 Dictyoste
31	22.4	70.0	165509	2	AC009886	AC009886 Homo sapi
C 32	22.4	70.0	168614	9	AL357672	AL357672 Human DNA
C 33	22.4	70.0	170672	2	AC113127	AC113127 Mus muscu
C 34	22.4	70.0	176968	2	AC129250	AC129250 Rattus no
C 35	22.4	70.0	179779	2	AC113106	AC113106 Mus muscu
C 36	22.4	70.0	180279	2	AC109717	AC109717 Rattus no
C 37	22.4	70.0	181835	9	AC026124	AC026124 Homo sapi
C 38	22.4	70.0	186681	9	AC068643	AC068643 Homo sapi
C 39	22.4	70.0	199806	2	AC025580	AC025580 Homo sapi
C 40	22.4	70.0	202490	2	AC121884	AC121884 Mus muscu
C 41	22.4	70.0	271144	9	HSXDPB	AL590763 Homo sapi
C 42	22.2	69.4	195761	2	AC113204	AC113204 Mus muscu
C 43	22	68.8	87767	2	AC014497	AC014497 Drosophi
C 44	22	68.8	94588	2	AC116196	AC116196 Rattus no
45	22	68.8	150587	3	AC007549	AC007549 Drosophi

ALIGNMENTS

RESULT 1
BTTN4430
LOCUS
DEFINITION
ACCESSION
X07651
VERSION
X07651.1
KEYWORDS
SOURCE
ORGANISM
Bacillus thuringiensis
REFERENCE
AUTHORS
Mahillon,J. and Lereclus,D.
TITLE
Structural and functional analysis of Tn4430: identification of an
integrase-like protein involved in the co-integrate-resolution

4149 bp DNA linear BCT 12-SEP-1993
Bacillus thuringiensis transposon Tn4430.
GI:40347
tnpA gene; transposase; transposon.
Bacillus thuringiensis.
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1. (bases 1 to 4149)
Mahillon,J. and Lereclus,D.
Structural and functional analysis of Tn4430: identification of an
integrase-like protein involved in the co-integrate-resolution

process	EMBO J. 7 (5), 1515-1526 (1988)
JOURNAL	88312602
MEDLINE	2842151
PUBMED	
COMMENT	Data kindly reviewed (03-APR-1989) by Lereclus D.
FEATURES	Location/Qualifiers
source	1..4149
	/organism="Bacillus thuringiensis"
	/strain="H1.1"
	/db_xref="taxon:1428"
	/map="plasmid pGI2"
	/clone="pGI200"
	/clone_lib="pBR322"
repeat_unit	1..38
	/note="inverted repeat A"
CDS	247..1101
	/note="put. resolvase Tnp I (AA 1 - 284)"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA30491.1"
	/db_xref="GI:40348"
	/translation="MDVAKQFSSYLKQENKTNVQGYTSGIRQYIKWFEGSYDRKLT
	KLYRONILEYISLVKQVLMNKAISNHKISSIAKFNELIQKQSQDQVILKTMKV
	QTVASPTQIVELDVKKFLOSLVDNKNRNYATLLAYTVGRISSEALSIMKDFNLQ
	TGECIRSGKGGKQRIVLNLSKVLSAIKDYLDIKRTYSTAHESPYLFTSKRKLDR
	VNRTFKSYNVITPHQLRHFFCTNAIEKGFSEHVAQAGHSNIHTTLLTNPQLQ
	LKNKMWLL"
CDS	1120..4083
	/note="transposase Tnp A (AA 1 - 987)"
	/codon_start=1
	/transl_table=11
	/protein_id="CAA30492.1"
	/db_xref="GI:40349"
	/db_xref="SWISS-PROT:P10021"
	/translation="MGVKQLSEAGNELMDLSRLTEWDLVTRFSSKDLHLILKHR
	RGYNRGLFALRLVIRYPCWSTLTKYDIPQYVAVYTSRLRIPPEFLVYAKRGNTLW
	EHLGETIRYEQNFSEYKELTLOFVQAMDNNTLYLIEITISLTKTKVILPAM
	VYEDIVWEAKQADQKQVYSLHDGLVQEQDQDALLPTINGKSPILAWLKDVPAPQ
	SPESFLKVIDRQFQVQIGLITDTTKINTNRLRLQARLGSKYEPVAFRENFVKRYSM
	LYSFLLEITQDLIDYAIETHDRMLMNLOTGKKEQDEIQOANGKLNKILQFIVCG
	TLFEAKETGDAPAAALDEVMNEMVSEAEKQSRPNLYDILLNTRYSYVRVYA
	PTLLSLHFRATKSGEPVQLQADTHLNETGKRVKPHGAPLHFVSNRWKQHVYDDDG
	NIWRHYELAAALTELNRHISRGDIFVSGRRHKAFDDYLIPIYDENWEVSNIPNGTAP
	LKREDYITDRINLNEHLWLSKNEKLEGVDSOGKLVHRLDRTGTPPEAKAFSKLL
	HSMPLRIKTLTDLIEVASWTGHDQFIHASTNOSPQEQEONIVLATLMAMGTNIGLTK
	MAEATPGISYROMANASQRMWDDAMVRAQSILVNFQKEKLSYWGDLTSSSDGMR
	LSTAVRSLHADSHPHYTGKGGTIIYFVSDQSAYHVKVTITNARDALHVLQGLLHHE
	TDLKTIEHTYDTAGYTDQFALTHLLGFRFAPRIRDLADTKLFSIPGGEYENVQALL
	KGKINVKLIKNEYEDIRRLAYSVQTKVSSALIMGKLSYARONKLAVALGEMGRIEK
	TLFTLDYISNKAVRRVQKGLNGKEAINALARTIFFQGEFRERALQDQLOQASALN
	IIINATSVNTVMYMEKAVEELKARGEFREDLMPYAWPLGWEHINFLGKFKFGLHDTG
	QMLNRLRIKEPYS"
repeat_unit	4112..4149
	/note="inverted repeat A"
BASE COUNT	1506 a 685 c 776 g 1182 t
ORIGIN	
Query Match	100.0%; Score 32; DB 1; Length 4149;
Best Local Similarity	100.0%; Pred. No. 0.39; 0; Indels 0; Gaps 0;
Matches	32; Conservative 0; Mismatches 0;
Qy	1 TAATACACACATTAATTAATGTTGTATTATTA 32
Db	113 TAATACACACATTAATTAATGTTGTATTATTA 144
RESULT 2	
LOCUS	BTPIG2XX
DEFINITION	Bacillus thuringiensis plasmid pGI2 with transposon Tn4430.
ACCESSION	X13481
VERSION	X13481.1 GI:3171732

IIINAISVNTVYMEKAVEELKARGEFDRLMPYAWPLGWEHINFLGEYKFEGLHDTG
OMNLRPLRIKEPFYS"

repeat_unit 4609..4646
/note="inverted repeat A"
repeat_region 4647..4651
/note="direct repeat 1"
CDS 4859..5473
/note="ORF 1"
/codon_start=1
/transl_table=1
/protein_id="CAA31836.1"
/db_xref="GI:3171733"
/db_xref="SWISS-PROT:P10022"
/translation="MTNIDFKHVKENARIRDFENEKEKFKODHNGINQEEVNOAMOV
LSKATGKEIFIGTKRPSQKVKFAQFIQDNWDYALENAFFTDEMLFLRIORPLOF
KSNICVNDIHSRNALPMOKOIADRLKTDKSKISIVNSLVQKGIYKANGHKPEGVK
ARTYALFNPNIIYSGERDNETTLKALFPMKSKSLFKFPPIALF"
6037..6849
/note="ORF 2"
/codon_start=1
/transl_table=1
/protein_id="CAA31837.1"
/db_xref="GI:3171734"
/db_xref="SWISS-PROT:P10023"
/translation="MKKLISFIAILFFICGFNLKAFAAEEIIDYQSLYNOAIOBQVL
DQNSVYNEWLNQKKEEMPIYQDGLKQGVFELPSYNEWLKLNNGYGOAPTGDILFD
DVTPRSGWGTFLKAGDIFITNATSSAGIVGHAALANGDNVILHMPGAGQNOQLSTS
NMOKYTAGSGWKVRLKDOTLARDVARYADRNFYSTGSAATKNVILDYGDIDHLYQ
KNPTVCSKLVLQALFYFGSGSRNVQAVSGIVTPYGLDITFTSAVRPSLVKY"
6896..7240
/note="ORF 3"
/codon_start=1
/transl_table=1
/protein_id="CAA31834.1"
/db_xref="GI:40321"
/db_xref="SWISS-PROT:P10024"
/translation="MMAIFISIVLVSEVLFNVRQAKSKIYNIARQIOBESQLKYID
FHKDFKMGYVLAVVEGENPDIIYEYSQDKKYNFQAYFNSEKAIRKKMMGGSLTE
IEMKKLYPPLQ"
7829..9166
/codon_start=1
/transl_table=1
/product="put. GI2 site-specific recombinase"
/protein_id="CAA31835.1"
/db_xref="GI:40322"
/db_xref="SWISS-PROT:P10025"
/translation="MNFPAIIMOKFQISDVQGTQKHQROGKSKSNLIDIDYSKQON
YDLNQKIRIESTIKQISERVKRPPRANSVVLSEFVVTASPDYMHSLSEOKRYF
ESSLDFIQKRYGKQNTLYAMVHDEATPHMHGYMPTEDNRLSAKDMFTREKILSLQ
QDFPLEMREKGFVDVDRGSEBKHLSPQAFKEKODLEVEQLSNVTKHLKTKVYVEE
NLOQNTNYIEKQNETLQIQOQLSLDKIKKKQKPEFTFRNOIPDKPVSMSYLREE
TKTEVTKLFGKPEITKKTGNIVTVREQRDMTEKNAAVIVKDYERLQKTDLYKE
NOSUREDKNLYEETIKGNLAKHSYKQNRLEEVNKLHEITGLKKAHIDLOMNIK
VLYQTKVFEQKFAKGLIKNELMDMGVQNPQEREHREIRSRQKQYDMER"

BASE COUNT 3574 a 1419 c 1788 g 2891 t

ORIGIN

Query Match 100.0%; Score 32; DB 1; Length 9672;

Best Local Similarity 100.0%; Pred. No. 0.32; Mismatches 0; Indels 0; Gaps 0;

Matches 32; Conservative 0;

QY 1 TAATACACACAAATATTAATGTTGTTGTTATTA 32

|||||

Db 610 TAATACACACAAATATTAATGTTGTTGTTATTA 641

|||||

RESULT 3

BTTN4430/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

BTNN4430

Bacillus

thuringiensis

transposon Tn4430.

x07651

GI:40347

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

plasmid; resolvase; tnpA gene; tnpI gene; transposase; transposon.

SOURCE
ORGANISM

Bacillus thuringiensis.
Bacillus thuringiensis
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus cereus group.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 4149)
Mahillon, J. and Lereclus, D.
Structural and functional analysis of Tn4430: identification of an
integrase-like protein involved in the co-integrate-resolution
process

JOURNAL
MEDLINE
PUBMED

EMBO J. 7 (5), 1515-1526 (1988)
88312602
2842151

COMMENT
FEATURES

Data kindly reviewed (03-APR-1989) by Lereclus D.
Location/Qualifiers
1..4149
/organism="Bacillus thuringiensis"
/strain="H1.1"
/db_xref="taxon:1428"
/map="plasmid pGI2"
/clone="pGI200"
/clone_lib="pBR322"
1..38
/note="inverted repeat A"
247..1101
/note="put. resolvase Tnp I (AA 1 - 284)"
/codon_start=1
/transl_table=1
/protein_id="CAA30491.1"
/db_xref="GI:40348"
/db_xref="SWISS-PROT:P10020"
/db_xref="GI:40349"
/note="transposase Tnp A (AA 1 - 987)"
/codon_start=1
/transl_table=1
/protein_id="CAA30492.1"
/db_xref="GI:40349"
/db_xref="SWISS-PROT:P10021"
/translation="MGVKLLSEARQNELMDLSRLTEWDLVTFRTESKHLHLILKUR
EGYNRLGALRLVLRYPGWSLTYKDIPOVAVVTSRLRIPPEEFLVYAKRGNTLW
YVIEDVWEAKQADQKQVSYSLHDGLVQEQDLDALLPTINGKSPWLKDVPAQ
SPESFLKVIDRQVQKIGLITDTTKINTNRLQRLARLQKGYEYAFRRFNEKRYSM
LVSLLEITQDIDYAIHDLRMNLTQKKEQDEIQOANGKLNKILQFTITVCG
TLIBAKETGDAFALDEVMSENVESVEAKOLSRPLNDYLDLNLTRYSYVRYA
PTLLRSLHFRATKSGEPVLOALDIHELNETGKRKYPHAGPLHFVSNRQKHVDDGG
NINHYELAALTELNRHISGDIIVSGSRHKKAFDDYLIPTDEWNEVSNIPNGLTAP
LKAEYITDRINRLNLEWLSKSEKLEGVDSIOGKLHVELDRGTPEEAKATSKLL
HSMPLRKLTLDLLEVASWTGFHQDFTASTNQSPDQEQEINIVLATLMAMGTNIGLTK
MAEATPGISYRQMANASQWMDAMVRAQSLVNFQKQKLSYVWGDTTSSDGMK
LSIAVRSLSHDSNPHYTGKGTIVRFVSDOLSVAHVKVITTNARDALHVDLGLLHHE
TDLKECHYTDYAGTDOVEALTHLGFAPRIEDLADTKLFSIPGCEEEVNVQALL
KCKINVKIKENYEDIRRLAYSQVOTGKVSALLMGKLSYARQNKALQDEMGRIEK
TLFTLDYISKNYRRVQKGLNKGEINALLRTIFFGQGRFERRFADQQLQASALN
IIINAISVNTVYMEKAVEELKARGEFDRLMPYAWPLGWEHINFLGEYKFEGLHDTG
OMNLRPLRIKEPFYS"
4112..4149
/note="inverted repeat A"
repeat_unit 1506 a 685 c 776 g 1182 t

repeat_unit

CDS

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

1120..4083

```

RESULT 4
BTPG12XX/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BTPG12XX
Bacillus thuringiensis plasmid pG12 with transposon Tn4430.
X13481.1 GI:3171732
plasmid; plasmid pG12; recombinase; resolvase; transposase;
transposon; unidentified reading frame.
Bacillus thuringiensis.
Bacillus thuringiensis.
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus; Bacillus
cereus group.
1 (bases 1 to 6999)
Mahillon, J. and Seurinck, J.
Complete nucleotide sequence of pG12, a Bacillus thuringiensis
plasmid containing Tn4430
Nucleic Acids Res. 16 (24), 11827-11828 (1988)
89098342
3211758
2
Mahillon, J.
Direct Submission
Submitted (04-NOV-1988) Mahillon J., Plant Genetics Systems, J
Plateaustaat 22, B-9000 Gent, Belgium
revised by [3]
3 (bases 1 to 9672)
Hoflack, L.
Direct Submission
Submitted (24-MAR-1998) Hoflack L., Plant Genetics Systems, J
Plateaustaat 22, B-9000 Gent, Belgium
On Jun 2, 1998 this sequence version replaced gi:40316.
FEATURES
Source
1..9672
Location/Qualifiers
/organism="Bacillus thuringiensis"
/strain="thuringiensis HI.1."
/db_xref="taxon:1428"
/plasmid="pG12"
493..4651
/transposon="Tn4430"
493..497
/feature="direct repeat 1"
498..535
/feature="inverted repeat A"
744..1598
/codon_start=1
/transl_table=11
/product="TnpI resolvase"
/protein_id="CAA31832.1"
/db_xref="GI:40317"
/db_xref="SWISS-PROT:P10020"
translation="MDVAKQPSYLKQENKNTENTVOGYTSIGROYIKWFEQSYDRKLT
KLYRNLLEYISYLNKVMNAKSTNNKRISSIAKFNFEFLIQSGQDDVILKTDIMKY
QTGYASPTQIVELDVKKFLQSVLEDDNNKNRYAIATPLLAYTVRISEALSIKMNDFNLQ
TGECIRSGKGKQRIVLNLSKVLSAIKDYLDIRKTYSTAHESPYLFISKRKRLDRT
VYNNRIFKSYNVITPQHLRHFFCTNAIEKGFSTIEHVAQAGHSNIHTTLLATNPQLQ
LNKNMELL"
1617..4580
/codon_start=1
/transl_table=11
/product="TnpA transposase"
/protein_id="CAA31833.1"
/db_xref="GI:40318"
/db_xref="SWISS-PROT:P10021"
translation="MGVKQLSEAGRNELMDLSRLTEWDLVTFRTFSKHDLLHLKHR
RGYNRLGALLRLVIRPGWSLLEYKDIPQYVAVXTVSRLEKIPPEFLVYAKRGNLW
EHLGIRTEYGYQNFSSYKETTLLQFLVQQAMNNNTLYLIEITISLTKTKVILPAM
YVIEDIVWEAKQADQKVYSILHDLGVQEKDQLDALLPTINGKSPFLAWLKDVPAQP
SPESFLKVIDQFQKIGLITDTTKINTNRLRLARIGSKVEPYAFRNEVKRYSM
LYSFLELLETDLIDYAIHEHRLMNLQTKGKKEDELOQANGKLNKILQFTVCG
TLIEAKETGDAFALDEVNSWENWSEAEAKQLSRPINDYLDLLNTRYSYVRYA
PTLLKSLHFRATKSGEPVLQALDTHIELNETGKRKVPHPGAPLHFVSNWQKHVYDDDG
NINHYELAAALTELNRHIRSGDIFVSGSRHHKAPDDYLIPIYDSEWNEVSNIPNGLTAP
LKAEDYITDRINLNEHLSKSEKLEGIQSLQHLVRLDRGTPPEAKAFSKLL
HMLPRIKLTLDLIEVASWTGHQFTHASTNQSPDQEQNIIVLTIATMAMGTNIGLTK
MAETPGISROMANASOMRYDDAMVRAQSILVNFQKEQLSSYWGDTTSSSDGMR
LSIAVRSIHADSNPHYGTGGTIVRFVSDOLSAAHVKVIITTNARDALHVLDDGLLHHE
TDLKIEHYDTAGTYTQVFTALHLLGFRFAPRIRDLATKLSIPGGEYENYVQALL
KGINVKLIKENYEDIRLAYSQVGTGKVSALIMGKLSYARQNKALATALGAEKRIEK
TLFTDLYISKNAVRVRQKNGKEATINALARTIFFQORGEFRERALQDOQLQASALN
IIINAISVNTVYMEKAVEELKARGEFRDLMPYAWPLGWEHINFLGKEYKPEGLHDTG
QMLRPLRIKEPEYS"
4609..4646
/feature="inverted repeat A"
4647..4651
/feature="direct repeat 1"
4859..5473
/feature="ORF 1"
/codon_start=1
/transl_table=11
/protein_id="CAA31836.1"
/db_xref="GI:3171733"
/db_xref="SWISS-PROT:P10022"
translation="MTNTIDFKHVEKNARIIDFENEKEFKQDHNGINQEEVNOAMOV
LSKATGKEIGTGRSPQSKVFAQFIQDNWDYALENAFTDEMLELRLIQRFLOF
KSNCTIVDIHSRNALPMSQKQIADRLKTDKRSIVNSLVOKGVIVKANGHKPEGVK
ARTYALFINPNIYSGERDNETTLKALFMNSKSLFKFPALF"
6037..6849
/feature="ORF 2"
/codon_start=1
/transl_table=11
/protein_id="CAA31837.1"
/db_xref="GI:3171734"
/db_xref="SWISS-PROT:P10023"
translation="MKLLISFIALLFFICGFLNKAFAAEIIDIYOSLYNOAIOBGLV
DONSYSNEWLKQNEKPEMPYQDLKQGVLEPLSYNEWLKLNNGYGAFTGDIFFLD
DYTPRSGWGTFLKAGDIFITNATSSAGIVGHAALANGDNVILHMPGAGQGNQLSTS
NMWKYTSAGSKWIKVYRUKDQTLARDVARYADRNFTYITGSAITNVIDIGDTHLYQ
KNPTYCSKLVFOALYFGSGSRNVQAVSGIVTPYGLIDTFTTSAYRPSLVKTY"
6896..7240
/feature="ORF 3"
/codon_start=1
/transl_table=11
/protein_id="CAA31834.1"
/db_xref="GI:40321"
/db_xref="SWISS-PROT:P10024"
translation="MMAIFISIVLVSPFLFNRVQAKSIYNIARQIOESQLKYID
FKDFKMGGYVLAVVVEGENPDIVVEYSYQDKVNFQAYFNSEAKIKKMMGGSGLTE
IEMKKLYPPLQ"
7829..9166
/codon_start=1
/transl_table=11
/product="put. G12 site-specific recombinase"
/protein_id="CAA31835.1"
/db_xref="GI:40322"
/db_xref="SWISS-PROT:P10025"
translation="MKNKFAIHMOKFQISDVQGTQKHQNRQKSKSNLIDIDYSKQNF
YDLNQQRIRYESTIKQIESRVRKPRANSVILSEFVYVATSPDYMHSLSLEOKRYF
ESSLDFIQKRYGKQNTLYAMVHDEATPHMIGVMPITEDNRLSAKDMFTKRLISLQ
ODEPLEMEKGFVDVRGSEKHLSPQAFKQKODLEVELOLNVKTHLTKVYVETH
NLOQOTTNYIEKQNETLQKIQOQFLSLDKKIKEKQEFETERNQIPDKPVSMLYLRRE
TKTEVTHLFGKPEITEKTKGNVYVTRQWRDMTEKVNAAVIVKDYERLQKTDLVKE
NQSLEDNKLYLEETIKGNLAKHSYKONRELEEVNKLHTEIGTLKAIRHQLQMNIAK
VLYQOTKRVKFEQKFAKRLIKNELMDKMGVDNQNFEREHTRISRQKGYDMEH"
BASE COUNT 3574 a 1419 c 1788 g 2891 t
ORIGIN
Query Match 80.0%; Score 25.6; DB 1; Length 9672;
Best Local Similarity 87.5%; Pred. No. 40;
Matches 28; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 TAATACACACAAATTAATATTGTTGTGTTATTA 32
|||||
DB 641 TAATACACACAAATTAATATTGTTGTGTTATTA 610
|||||

```



```

RESULT 5
AC027541/c
LOCUS
DEFINITION
  Homo sapiens chromosome 8 clone RP11-324F11 map 8, linear HTG 17-AUG-2002
  IN PROGRESS ***, 1 ordered piece.
AC027541.1 GI:22296720
VERSION
AC027541.1
KEYWORDS
  HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE
  Human.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 164303)
  Birren,B., Nusbaum,C. and Lander,E.
  Homo sapiens chromosome 8, clone RP11-324F11
  Unpublished
  2 (bases 1 to 164303)
  Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
  Anderson,S., Baldwin,J., Barna,N., Bastien,V., Bida,F.,
  Boguslavskiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
  Campopiano,A., Castele,A., Choepel,Y., Colangelo,M., Collins,S.,
  Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
  Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
  Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
  Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
  Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
  Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,
  Levine,R., Liu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
  McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
  Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V.,
  Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
  Pisanli,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
  Roy,A., Santos,R., Schauer,S., Severi,P., Spencer,B.,
  Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
  Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
  Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
  Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  3 (bases 1 to 164303)
  Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
  Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
  Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
  Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
  Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
  Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,
  Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
  Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-roh,K.,
  Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
  McCarthy,M., Meldrim,J., Meneus,L., Mihova,T., Mienga,V.,
  Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
  O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
  Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
  Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severi,P.,
  Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
  Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
  Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
  Zembek,L., Zimmer,A. and Zody,M.
  Direct Submission
  Submitted (17-AUG-2002) Whitehead Institute/MIT Center for Genome
  Research, 320 Charles Street, Cambridge, MA 02141, USA
  On Aug 17, 2002 this sequence version replaced gi:22123707.
  All repeats were identified using RepeatMasker:
  Smit, A.F.A. & Green, P. (1996-1997)
  http://ftp.genome.washington.edu/RM/RepeatMasker.html
  ----- Genome Center
  Center: Whitehead Institute/ MIT Center for Genome Research
  Center code: WIGR
  Web site: http://www-seq.wi.mit.edu
  Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L8782
Center clone name: 324_F_11
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 164303: contig of 164303 bp in length.
FEATURES
      source
      Location/Qualifiers
        1..164303
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /chromosome="8"
        /map="8"
        /clone="RP11-324F11"
        /clone_lib="RPCI-11 Human Male BAC"
BASE COUNT 48256 a 33368 c 34914 g 47765 t
ORIGIN
Query Match 76.9%; Score 24.6; DB 2; Length 164303;
Best Local Similarity 87.1%; Pred. No. 44;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AATACACACATATTAAATTTGTTCTGTTATTA 32
      ||||| ||||| ||||| ||||| |||||
Db 116845 AATACATACATATTAAAGTGTGTTTATTA 116815
      ||||| ||||| ||||| ||||| |||||

RESULT 6
AC100807
LOCUS
DEFINITION
  Homo sapiens chromosome 8 clone CTD-2534J5 map 8, linear HTG 20-AUG-2002
  PROGRESS ***, 2 unordered pieces.
AC100807.2 GI:22325314
VERSION
AC100807.2
KEYWORDS
  HTG: HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE
  Human.
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 176759)
  Birren,B., Nusbaum,C. and Lander,E.
  Homo sapiens chromosome 8, clone CTD-2534J5
  Unpublished
  2 (bases 1 to 176759)
  Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
  Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
  Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
  Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,A.,
  Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
  Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
  Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
  Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
  Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
  Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
  Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
  McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
  Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
  Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
  Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Rise,C., Rogov,P.,
  Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
  Seaman,S., Severi,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
  Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
  Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
  Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
  Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

sequence.
AL663040 GI:18855245
AL663040.6
HTG.
house mouse.
Mus musculus
1 (bases 1 to 81885)
Pearce,A.
Direct Submission
Submitted (16-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humques@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 21, 2002 this sequence version replaced gi:18375888.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/C-elegans/wormpep RP23-61M22 is
from the RPI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
IMPORTANT: This sequence is not the entire insert of clone
RP23-61M22 It may be shorter because we sequence overlapping
sections only once, except for a short overlap.
The true left end of clone RP23-82B4 is at 79886 in this sequence.
The true right end of clone RP23-206116 is at 2000 in this
sequence.

FEATURES
source
1..81885
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-61M22"
/clone_lib="RPI-23"
BASE COUNT 26614 a 14835 c 14385 g 26051 t
ORIGIN
Query Match 71.9%; Score 23; DB 10; Length 81885;
Best Local Similarity 83.9%; Pred. No. 1.7e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TAATACAACAATATTAATTTGGTTGTATT 31
||||| ||||||| ||||||| |||||||
Db 66034 TAATACAACAATATTAATACAATTTGTTATT 66064
RESULT 8
AC083886/c
LOCUS
DEFINITION Homo sapiens 3 BAC RP11-364F11 (Roswell Park Cancer Institute Human
BAC Library) complete sequence.
ACCESSION AC083886
VERSION AC083886.19 GI:21956512
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 176759)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukghalter,B.,
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karacas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Meneus,L., Mihova,P., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Testaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 20, 2002 this sequence version replaced gi:17048177.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21470
Center clone name: 2534_J_5

* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 117162: contig of 117162 bp in length
* 117163 117262: gap of 100 bp
* 117263 176759: contig of 59497 bp in length.
FEATURES
Location/Qualifiers
1..176759
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="CTD-2534J5"
/clone_lib="CITD2 Human BAC"
BASE COUNT 52709 a 37301 c 35610 g 51038 t 101 others
ORIGIN
Query Match 76.9%; Score 24.6; DB 2; Length 176759;
Best Local Similarity 87.1%; Pred. No. 43;
Matches 27; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 AATACAACAATATTAATTTGGTTGTATTA 32
||||| ||||||| ||||||| |||||||
Db 147842 AATACAACAATATTAAGTGTGTTTATTA 147872
RESULT 7
AL663040
LOCUS
DEFINITION Mouse DNA sequence from clone RP23-61M22 on chromosome 11, complete

REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1 (bases 1 to 94535)
 Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-Osman, F.R., Allen, C.,
 Alsbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T.,
 Barbarella, J., Benton, J., Blumage, K., Blankenburg, K., Bonnin, D.,
 Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P.,
 Buhay, P., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
 Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
 Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
 Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
 Davila, M.L., Davis, C., Davy-Garroll, L., Dederich, D.A.,
 Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
 Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
 Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
 Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
 Gabisi, J., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, W., Gunaratne, P., Haley, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X.,
 Hernandez, J., Hernandez, O., Hodgson, A., Hogue, M., Holloway, C.,
 Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,
 Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
 Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
 Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, B.C.,
 Lewis, L., Li, J., Li, Z., Lichtarge, O., Lied, C., Liu, J., Liu, W.,
 Lounsged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
 Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
 Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M.,
 Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
 Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D.,
 Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,
 Nickerson, E., Nwokenwo, S., Ogih, M., Okwuonu, G., Oragunye, N.,
 Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,
 Pickens, R., Prius, E., Pu, L.D., Quiles, M., Ren, Y., Rives, M.,
 Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S.,
 Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E.,
 Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,
 Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tausey, J., Taylor, C.,
 Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmami, K., Vasquez, L.,
 Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,
 Warren, R., Washington, C., Watlington, S., Williams, G.,
 Williamson, A., Wlecezyk, R., Woodson, S., Worley, K., Wu, C., Wu, Y.,
 Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and
 Gibbs, R.

Direct Submission
 Unpublished
 2 (bases 1 to 94535)
 Worley, K.C.

Direct Submission
 Submitted (05-OCT-2000) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 94535)
 Worley, K.C.

Direct Submission
 Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

4 (bases 1 to 94535)
 Worley, K.C.

Direct Submission
 Submitted (26-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE
AUTHORS
JOURNAL

REFERENCE
AUTHORS
JOURNAL

REFERENCE
AUTHORS
JOURNAL

COMMENT
 On Jul 26, 2002 this sequence version replaced gi:21952622.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
 gc-help@bcm.tmc.edu

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annot> ation.html.

FEATURES	Source	Location/Qualifiers
		1. .94535
		/organism="Homo sapiens"
		/db_xref="taxon:9606"
		/chromosome="3"
		/clone="RP11-364F11"
		1145..1277
repeat_region		/rpt_family="MIR"
repeat_region		1365..1393
repeat_region		/rpt_family="AT rich"
repeat_region		complement(3179..3270)
repeat_region		/rpt_family="L2"
repeat_region		4425..4494
repeat_region		/rpt_family="(TC)n"
repeat_region		5662..5848
repeat_region		/rpt_family="MIR"
repeat_region		5911..6195
repeat_region		/rpt_family="AluSc"
repeat_region		complement(6661..6806)
repeat_region		/rpt_family="LTR33"
repeat_region		complement(6904..7076)
repeat_region		/rpt_family="LTR33"
repeat_region		7269..7372
repeat_region		/rpt_family="L2"
repeat_region		complement(9820..9931)
repeat_region		/rpt_family="L1MA8"
repeat_region		9932..10236
repeat_region		/rpt_family="AluSc"
repeat_region		complement(10237..10651)
repeat_region		/rpt_family="L1MA8"
repeat_region		11144..11681
repeat_region		/rpt_family="L2"
repeat_region		complement(12092..12167)
repeat_region		/rpt_family="MIR"
repeat_region		12470..12754
repeat_region		/rpt_family="L1MA1"
repeat_region		complement(14419..14672)
repeat_region		/rpt_family="MIR"
repeat_region		complement(14920..15017)
repeat_region		/rpt_family="MER5A"
repeat_region		15660..15729
repeat_region		/rpt_family="MIR"
repeat_region		16326..16354
repeat_region		/rpt_family="(CA)n"
repeat_region		complement(16742..16969)
repeat_region		/rpt_family="MIR"

Thu Feb 13 12:34:53 2003

```

repeat_region 17622..17672
/rpt_family="L2"
repeat_region 18330..18481
/rpt_family="L2"
STS 18834..19127
/standard_name="SHGC-106103"
repeat_region 20446..20565
/rpt_family="MIR"
repeat_region 20616..20737
/rpt_family="MIR"
repeat_region 21408..21443
/rpt_family="(CA)n"
repeat_region 22684..22746
/rpt_family="GA-rich"
repeat_region 24196..24261
/rpt_family="MIR"
repeat_region 25871..25939
/rpt_family="(TA)n"
STS 28818..29072
/standard_name="D3S3421"
repeat_region 29450..29477
/rpt_family="(T)n"
STS 29509..29817
/standard_name="RH121121"
STS 30562..30711
/standard_name="SHGC-33976"
STS 30683..30820
/standard_name="A009Y32"
repeat_region 32501..32646
/rpt_family="MIR"
repeat_region complement(34074..34228)
/rpt_family="MIR"
repeat_region 35022..35138
/rpt_family="(TATAA)n"
repeat_region 35917..35972
/rpt_family="L2"
repeat_region 38774..38822
/rpt_family="L2"
repeat_region 43021..43124
/rpt_family="MIR"
repeat_region 45239..45261
/rpt_family="(TTTC)n"
repeat_region 46378..46691
/rpt_family="AluY"
repeat_region 47134..47154

Query Match 71.9%; Score 23; DB 9; Length 94535;
Best Local Similarity 83.9%; Pred. No. 1.7e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AATACACACAAATATTGTTGTTGTTATTA 32
|||||
DB 56960 AATATACACAAATATTACTTTTGTGATGA 56930

RESULT 9
AC003028
LOCUS AC003028 115359 bp DNA linear PLN 11-MAR-2002
DEFINITION Arabidopsis thaliana chromosome 2 clone F16M14 map ve018, complete
sequence.
ACCESSION AC003028
VERSION AC003028.3 GI:20196982
KEYWORDS HTG.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 115359)
Rounsley,S.D., Kaul,S., Lin,X., Ketchum,K.A., Crosby,M.L.,
Brandon,R.C., Sykes,S.M., Mason,T.M., Kerlavage,A.R., Adams,M.D.,
Somerville,C.R. and Venter,J.C.
REFERENCE
AUTHORS Unpublished
JOURNAL

```

```
/rpt_family="(GA)n"
join(13026..13450,14052..14502)
/gene="At2g38090"
/codon_start=1
/product="putative MYB family transcription factor"
/db_xref="GI:3335378"
/translation="MSPATYLFETSNLFOENRGTKWTAEENKTFENALAFYDKTDPDR
WSRVAMLPKGTIVGDIVIKOYRELEDDVSDIEAGLIPICYASDSEFTLDWGYDGASGN
NGFNMGYVFSAGGKRGSAARTAEHKKGVPTTEEHKQFLMGLKKYIGKGDNRNTA
RNFVTTRPTQVTOVSHAQKYFFIOVNGGKDKRRSSLIHDTITVINIPDSDAAADNATAN
APCSPSVGGNGRTESEWEGQLYDEATAAFYNQNAFSETLLGMSSTPYMAKLOEQSF
LNASFESINATLQM"
complement(13560..13621)
/rpt_family="(CATATA)n"
complement(14824..14855)
/rpt_family="AT-rich"
complement(15429..15450)
/rpt_family="(A)n"
complement(16211..17955)
/gene="At2g38100"
/note="F16M14.3"
complement(join(<16211..17060,17144..17637,17734..>17955))
/gene="At2g38100"
complement(join(16211..17060,17144..17637,17734..17955))
/gene="At2g38100"
/codon_start=1
/product="putative peptide/amino acid transporter"
/protein_id="AAC27159.1"
/db_xref="GI:3335358"
/translation="MSVLSWAFVAVFTLWMLYLTNEMKLFKFTDAAAIIVNVFAGVS
AIGHLGQFLVDAFIGHFWMCLTSLAFSGFGFLAISASPIILSGNGKGLFYVALTY
ISVGIFRSGSLGVFTDQLEDGRNKNPAKLVSVFIGNVGNFVFLAAIAIMPQISLP
WFVFTIPSGCEVLAFLIFISGACSYKRVKPGSGPLTTFVRFMASSKMSCAYSNNS
SQLYKAECDQDKPHRTSLRYLDRAAMILQTESLEOQRNWKLCVTEVEQTKSVI
RTVPLFATISLISGTVSLGNTFLEOAHNMDSKFSWNLPLPLLLFSEARLGSREL
CVMAAKRHAIIDFEPKPKTKTPYGPVSIILISFCCSIAAHVESRKLKVVSTOGLLHA
TVMSVFWLLPQVILGSLTIGVYIENSFALXLEETVPEELSQYMLLVNVCVGGVMSN
IALVLSVSGSKWFQDITNKSRVDNYWVITVFCMFLNLLFYIVTRYTVCKNKDG
ATQENDRRITASV"
complement(20543..23154)
/gene="At2g38110"
/note="F16M14.4"
complement(join(<20543..21412,22456..>23154))
/gene="At2g38110"
complement(join(20543..21412,22456..23091))
/gene="At2g38110"
/codon_start=1
/product="unknown protein"
/protein_id="AAC27160.1"
/db_xref="GI:3335359"
/translation="MGAQKRRRFEQISKCDVKDRSNHTVAAADLDGTLILSRSAFPYY
FLVALRAGSLRLALILVSVPPVYLYLTISLAINVEVFTFAGLKIIRDVLEVVRYS
VLPRTFAEDVRPTWIRFNTFKRYIITASPRIMVEPVFKTELGVDKVLDELYVRS
GRFTFRKPLGLVGQYKRDVLRVFGGLASDLPLDGLDSTKTDHDEMSICKEGYMP
RTKCEPLKRLPLSIIFHEGLRVORPTPLVALLFLPLPGVFLSIIRVYINILPE
RTARYNYKLTGLKLVNCGHPPPKPGQCHLLVNCNHRVLDPVYVALGKISCVT
YSISKELSIPKVALPROEKDAANKRLLEEGDLVICPEGTTCREPFLRFESAL
FAELTDRIVPVAINTKQSFNGTITGKLLDPYAFMNPRTYTEITFLKQIPAEITC
KGGKSPFVANTQIRVLGTLGTFECTNFKDKYAMLAGTDRGVPVKRKT"
23106..23191
/rpt_family="(GA)n"
23124..23191
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23124..23189
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23124..23183
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23128..23191
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23129..23193
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23131..23191
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23133..23193
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23135..23191
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23138..23193
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
23139..23191
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A A.thaliana
microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
74A"
29497..29526
Query Match 71.9%; Score 23; DB 8; Length 115359;
Best Local Similarity 83.9%; Pred. No. 1.6e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TAATACACACAAATATTAAATGTGTGTATT 31
||||| ||| ||||| ||||| ||||| |||||
Db 36293 TAATAAATAAATATTAAATGTGTATT 36323
RESULT 10
AL161631/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-70K10 on chromosome 9, complete
sequence.
ACCESSION AL161631
VERSION AL161631.20 GI:16972803
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 119819)
AUTHORS Hammond,S.
TITLE Direct Submission
JOURNAL
COMMENT
Submitted (15-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Nov 16, 2001 this sequence version replaced gi:15072563.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr-9
RP11-70K10 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see

http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
IMPORTANT: This sequence is not the entire insert of clone RP11-70K10 It may be shorter because we sequence overlapping sections only once, except for a short overlap.
The true left end of clone RP11-35N6 is at 117820 in this sequence.
The true right end of clone RP11-80H12 is at 2000 in this sequence.

FEATURES
source
1. .119819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="9"
/clone="RP11-70K10"
/clone_lib="RPCI-11.1"
BASE COUNT 36242 a 22548 c 23464 g 37565 t
ORIGIN

Query Match 71.9%; Score 23; DB 9; Length 119819;
Best Local Similarity 83.9%; Pred. No. 1.6e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TAATACACACAAATTAATTGTTGTATT 31
||||| ||||||| ||||| ||||| |||||
DB 71981 TAATAAACACAAATATTCATTTTATTATATT 71951

RESULT 11
AC017792/c
LOCUS
DEFINITION
AC017792 123579 bp DNA linear HTG 09-DEC-1999
Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
AC017792
AC017792.1 GI:6553398
HTG: HTGS_PHASE2.
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 123579)
Adams,M. and Venter,J.C.

Direct Submission
Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
This sequence was identified as CDM:10212105 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
Location/Qualifiers
1. .123579
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 35860 a 25889 c 26036 g 35734 t

ORIGIN

Query Match 71.9%; Score 23; DB 2; Length 123579;
Best Local Similarity 83.9%; Pred. No. 1.6e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2 RATAACACAAATTAATTGTTGTATT 32
||||| ||||||| ||||||| ||||| |||||
DB 36974 AACAAACACATATATTGTTATTATTTA 36944

RESULT 12
AC011096/c
LOCUS
DEFINITION
AC011096 145734 bp DNA linear HTG 28-SEP-2000
Homo sapiens clone RP11-2D21, WORKING DRAFT SEQUENCE, 4 unordered pieces.

AC011096
AC011096.3 GI:10334908
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens.

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 145734)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens, clone RP11-2D21
Unpublished
2 (bases 1 to 145734)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boukhalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeArelano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,K., Gage,D., Horton,L., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Headford,A., Klein,J., Leland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lechoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,K., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Direct Submission
Submitted (01-OCT-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 28, 2000 this sequence version replaced gi:7717103.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L1511
Center clone name: 2.D-21
----- Summary Statistics
Sequencing vector: M13; M7815; 93% of reads
Sequencing vector: Plasmid; n/a; 7% of reads
Chemistry: Dye-primer-amerisham; 3% of reads
Chemistry: Dye-terminator Big Dye; 97% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 144340 bases at least Q40
Consensus quality: 145024 bases at least Q30
Consensus quality: 145305 bases at least Q20
Insert size: 150000; agarose-1p
Insert size: 145434; sum-of-contigs
Quality coverage: 7.7 in Q20 bases; a.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* As soon as it is available and the accession number will
* be preserved.
*
* 1 41082: contig of 41082 bp in length
* 41083 41182: gap of 100 bp
* 41183 55282: contig of 14100 bp in length
* 55283 55382: gap of 100 bp
* 55383 111365: contig of 55983 bp in length
* 111366 111468: gap of 100 bp
* 111466 145734: contig of 34269 bp in length.
FEATURES
    source
        1..145734
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="RP11-2D21"
            /clone_lib="RPC1-11 Human Male BAC"
        1..41082
            /note="assembly_fragment"
            clone_end:SP6
            vector_side:left
        41183..55282
            /note="assembly_fragment"
        55383..111365
            /note="assembly_fragment"
        111466..145734
            /note="assembly_fragment"
            clone_end:T7
            vector_side:right
BASE COUNT 44971 a 27856 c 27629 g 44978 t 300 others
ORIGIN
Query Match 71.9%; Score 23; DB 2: Length 145734;
Best Local Similarity 83.9%; Pred No. 1.5e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TAATACACACAATATTATTCGTGTGTTATT 31
    ||||| ||||| ||||| ||||| |||||
Db 59434 TAATAAACACAATATTATTCATTTATTATTATTT 59404

RESULT 13
AC084188 151535 bp DNA linear HTG 15-JAN-2001
LOCUS Homo sapiens clone RP11-240G10, WORKING DRAFT SEQUENCE, 27
DEFINITION unordered pieces.
AC084188
VERSION AC084188.2 GI:12229375
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 151535)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 151535)
    Anderson,S., Barna,N., Bastien,V., Bedalov,F., Boguslavsky,L.,
    Boukhalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
    Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
    Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
    Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
    Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
    Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,
    Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
    Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,
    McPheeters,R., Melgrim,J., Meneus,L., Mihova,T., Mlenga,V.,
    Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
    O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
    Pierre,N., Pisanic,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
    Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
    Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
    Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
    Tirrelli,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,
    Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
    Zimmer,A. and Zody,M.
Direct Submission
Submitted (14-OCT-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 15, 2001 this sequence version replaced gi:10801416.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L11461
Center clone name: 240_G_10
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 142103 bases at least Q40
Consensus quality: 146385 bases at least Q30
Consensus quality: 148123 bases at least Q20
Insert size: 157000; agarose-fp
Quality coverage: 4.2 in Q20 bases; agarose-fp
Quality coverage: 4.4 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 27 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 241: contig of 241 bp in length
* 242 341: gap of 100 bp
* 342 1664: contig of 1323 bp in length
* 1665 1764: gap of 100 bp
* 1765 3932: contig of 2168 bp in length
* 3933 4032: gap of 100 bp
* 4033 5773: contig of 1741 bp in length
* 5774 5873: gap of 100 bp
* 5874 7178: contig of 1305 bp in length
* 7179 7278: gap of 100 bp
* 7279 8531: contig of 1253 bp in length
* 8532 8631: gap of 100 bp
* 8632 10607: contig of 1976 bp in length
* 10608 10707: gap of 100 bp
* 10708 33534: contig of 22827 bp in length
* 33535 33634: gap of 100 bp
* 33635 35523: contig of 1889 bp in length
* 35524 35623: gap of 100 bp
* 35624 38893: contig of 3270 bp in length
* 38894 38993: gap of 100 bp
* 38994 41672: contig of 2679 bp in length
* 41673 41772: gap of 100 bp
* 41773 44425: contig of 2653 bp in length
* 44426 44525: gap of 100 bp
* 44526 47448: contig of 2923 bp in length
* 47449 47548: gap of 100 bp
* 47549 52274: contig of 4726 bp in length
* 52275 52374: gap of 100 bp
* 52375 55715: contig of 3341 bp in length
* 55716 55815: gap of 100 bp
* 55816 58645: contig of 2830 bp in length
* 58646 58745: gap of 100 bp
* 58746 61720: contig of 2975 bp in length

```

```

* 61721 61820: gap of 100 bp
* 61821 66287: contig of 4467 bp in length
* 6288 66387: gap of 100 bp
* 66388 71117: contig of 4730 bp in length
* 71118 71217: gap of 100 bp
* 71218 77151: contig of 5934 bp in length
* 77152 77251: gap of 100 bp
* 77252 87532: contig of 10281 bp in length
* 87533 87632: gap of 100 bp
* 87633 94594: contig of 6962 bp in length
* 94595 94694: gap of 100 bp
* 94695 103008: contig of 8314 bp in length
* 103009 103108: gap of 100 bp
* 103109 113634: contig of 10526 bp in length
* 113635 113734: gap of 100 bp
* 113735 128598: contig of 14864 bp in length
* 128599 128698: gap of 100 bp
* 128699 149705: contig of 21007 bp in length
* 149706 149805: gap of 100 bp
* 149806 151535: contig of 1730 bp in length.

```

FEATURES

source

```

1. .151535
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-240G10"
/clone_lib="RPC1-11 Human Male BAC"

```

misc_feature

```

1. 241
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
342. .1664

```

misc_feature

```

/note="assembly_fragment"

```

misc_feature

```

1765. 3932
/note="assembly_fragment"

```

misc_feature

```

4033. .5773
/note="assembly_fragment"

```

misc_feature

```

5874. .7178
/note="assembly_fragment"

```

misc_feature

```

7279. 8531
/note="assembly_fragment"

```

misc_feature

```

8632. .10607
/note="assembly_fragment"

```

misc_feature

```

10708. .33534
/note="assembly_fragment"

```

misc_feature

```

33635. .35523
/note="assembly_fragment"

```

misc_feature

```

35624. .38893
/note="assembly_fragment"

```

misc_feature

```

38994. .41672
/note="assembly_fragment"

```

misc_feature

```

41773. .44425
/note="assembly_fragment"

```

misc_feature

```

44526. .47448
/note="assembly_fragment"

```

misc_feature

```

47549. .52274
/note="assembly_fragment"

```

misc_feature

```

52375. .55715
/note="assembly_fragment"

```

misc_feature

```

55816. .58645
/note="assembly_fragment"

```

misc_feature

```

58746. .61720
/note="assembly_fragment"

```

misc_feature

```

61821. .66287
/note="assembly_fragment"

```

misc_feature

```

66388. 71117
/note="assembly_fragment"

```

misc_feature

```

71218. .77151
/note="assembly_fragment"

```

misc_feature

```

77252. .87532
/note="assembly_fragment"

```

misc_feature

```

87633. .94594
/note="assembly_fragment"

```

misc_feature

```

94695. .103008

```

```

/note="assembly_fragment"
103109. .113634
/note="assembly_fragment"
113735. .128598
/note="assembly_fragment"
128699. .149705
/note="assembly_fragment"
149806. .151535
/note="assembly_fragment"
clone_end:T7
vector_side:right"
BASE COUNT 44763 a 29429 c 29282 g 45460 t 2601 others
ORIGIN

```

```

Query Match 71.9%; Score 23; DB 2; Length 151535;
Best Local Similarity 83.9%; Pred. No. 1.5e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

```

QY 2 AATACACACAATATTAATGTGTGTGATTATTA 32
||||| ||||| ||||| || ||||| ||

```

```

Db 59894 AATATAACACAATATTACTTTTTTTTGTATGA 59924

```

RESULT 14

AC008254/c

LOCUS

DEFINITION

AC008254

Drosophila melanogaster, chromosome 3R, region 85C-85C, BAC clone

AC008254.7

GI:12957659

HTG.

AC008254.7

Drosophila melanogaster.

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 164982)

REFERENCE

AUTHORS

Celnikier,S.E., Adams,M.D., Kronmiller,B., Tyler,D., Wan,K.H.,

Holt,R.A., Evans,C.A., Gocayne,J.D., Ananatides,P.G., Brandon,R.C.,

Rogers,Y., An,H., Baldwin,D., Bazon,J., Beeson,K.Y., Busam,D.A.,

Carlson,J.W., Center,A., Champe,M., Davenport,L.B., Dietz,S.M.,

Dodson,K., Dorsett,V., Doup,L.E., Doyle,C., Dresnek,D., Farfan,D.,

Ferreira,S., Frise,E., Galle,R.F., Garg,N.S., George,R.A.,

Gonzalez,M., Houch,J., Hoskins,R.A., Hostin,D., Howland,T.J.,

Ibegwan,C., Jalali,M., Kruse,D., Li,P., Mattei,B., Moshrefi,A.,

McIntosh,T.C., Moy,M., Murphy,B., Nelson,C., Nelson,K.A., Nunoo,J.,

Pacleb,J., Paragas,V., Park,S., Patel,S., Pfeiffer,B.,

Phouanavong,S., Pittman,G.S., Puri,V., Richards,S., Scheeler,F.,

Stapleton,M., Strong,R., Svirska,R., Tector,C., Williams,S.M.,

Zaveri,J.S., Smith,H.O., Rubin,G.M. and Venter,J.C. 85C-85C

Sequencing of Drosophila chromosome 3R, region 85C-85C

Unpublished

2 (bases 1 to 164982)

TITLE

JOURNAL

REFERENCE

AUTHORS

Celnikier,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G.,

Butenhoff,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L.,

Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L.,

Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L.,

Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P.,

Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S.,

Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Snir,E.,

Svirska,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and

Rubin,G.M.

Direct Submission

Submitted (30-JUL-1999) Drosophila Genome Center, Lawrence Berkeley

Laboratory, MS 64-121, Berkeley, CA 94720, USA

On Feb 17, 2001 this sequence version replaced gi:70181775.

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory, MS 64-121

Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome

shotgun and from subclones of this BAC and its neighboring clones.

For further information about this sequence, including its location

and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdgpe@fruitfly.berkeley.edu.

FEATURES

Source

```
1. .164982
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="85C-85C"
/clone="BACR02G22 (D833)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"
```

BASE COUNT 50436 a 33730 c 33344 g 47472 t

ORIGIN

```
Query Match 71.9%; Score 23; DB 3; Length 164982;
Best Local Similarity 83.9%; Pred. No. 1.5e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

QY 2 AATACACACAATATTAATTGTTGTTATTA 32
 |||||
 Db 91745 AACAAAGACAATATTAATTGTTATTTA 91715

RESULT 15

AC009258/c

LOCUS

DEFINITION AC009258 176036 bp DNA linear INV 17-FEB-2001
 Drosophila melanogaster, chromosome 3R, region 85C-85C, BAC clone
 BACR09F04, complete sequence.

ACCESSION

AC009258 AC009258.3 GI:12957665

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

1 (bases 1 to 176036)
 Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
 Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,
 Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,
 Carlson, J.W., Center, A., Champe, M., Davenport, L.B., Dietz, S.M.,
 Dodson, K., Dorsett, V., Doup, L.E., Doyle, C., Dresnek, D., Farfan, D.,
 Ferrier, S., Fris, E., Galle, R.F., Garg, N.S., George, R.A.,
 Gonzalez, M., Houch, J., Hoskins, R.A., Hostin, D., Howland, T.J.,
 Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A.,
 McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,
 Pacle, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Scheeler, F.,
 Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,
 Stapleton, M., Strong, R., Svirskas, R., Tector, C., Williams, S.M.,
 Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
 Sequencing of Drosophila chromosome 3R, region 85C-85C
 Unpublished

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 176036)
 Celniker, S.E., Aghayani, A., Arcaina, T.T., Baxter, E., Blazej, R.G.,
 Butenhoff, C., Champe, M., Chavez, C., Chew, M., Ciesiolka, L.,
 Doyle, C.M., Farfan, D.E., Galle, R., George, R.A., Harris, N.L.,
 Hoskins, R.A., Houston, K.A., Hummasti, S.R., Karra, K., Kearney, L.,
 Kim, E., Lee, B., Lewis, S., Li, P., Lomotan, M.A., Mazda, P.,
 Moshrefi, A.R., Moshrefi, M., Nixon, K., Pacle, J.M., Park, S.,
 Pfeiffer, B., Poon, L., Sequeira, A., Sethi, H., Snir, E.,
 Svirskas, R.R., Wan, K.H., Weinburg, T., Zhang, R., Zierán, L.L. and
 Rubin, G.M.

Direct Submission

TITLE

JOURNAL

COMMENT

Submitted (09-AUG-1999) Drosophila Genome Center, Lawrence Berkeley
 Laboratory, MS 64-121, Berkeley, CA 94720, USA
 On Feb 17, 2001 this sequence version replaced gi:5734650.
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory, MS 64-121
 Berkeley, CA 94720

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (<http://www.fruitfly.org/sequence/>) or send email to bdgpe@fruitfly.berkeley.edu.

FEATURES

source

```
1. .176036
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="85C-85C"
/clone="BACR09F04 (D967)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACe3.6)"
```

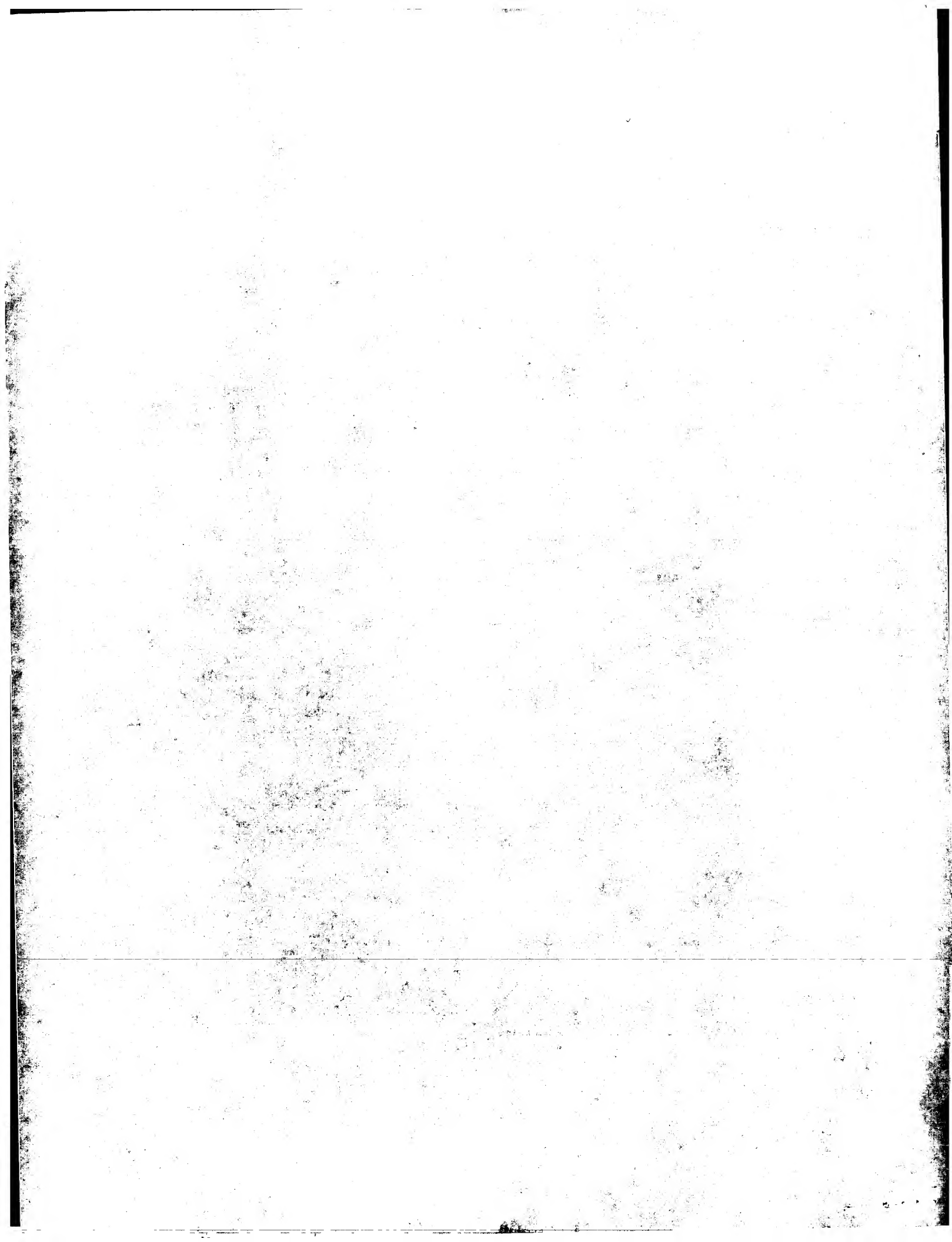
BASE COUNT 49927 a 37468 c 38131 g 50510 t

ORIGIN

```
Query Match 71.9%; Score 23; DB 3; Length 176036;
Best Local Similarity 83.9%; Pred. No. 1.4e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

QY 2 AATACACACAATATTAATTGTTGTTATTA 32
 |||||
 Db 12907 AACAAAGACAATATTAATTGTTATTTA 12877

Search completed: February 13, 2003, 02:44:17
 Job time : 3510 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2003, 22:30:55 ; Search time 291 Seconds
(without alignments)
247.642 Million cell updates/sec

Title: US-09-895-435-3

Perfect score: 32

Sequence: 1 taatacaacacataattgtgtgtatta 32

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125998159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1980.DAT:*
- 2: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1981.DAT:*
- 3: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1982.DAT:*
- 4: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1983.DAT:*
- 5: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1984.DAT:*
- 6: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1985.DAT:*
- 7: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1986.DAT:*
- 8: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1987.DAT:*
- 9: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1988.DAT:*
- 10: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1989.DAT:*
- 11: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1990.DAT:*
- 12: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1991.DAT:*
- 13: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1992.DAT:*
- 14: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1993.DAT:*
- 15: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1994.DAT:*
- 16: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1995.DAT:*
- 17: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1996.DAT:*
- 18: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1997.DAT:*
- 19: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1998.DAT:*
- 20: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA1999.DAT:*
- 21: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA2000.DAT:*
- 22: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA2001A.DAT:*
- 23: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA2001B.DAT:*
- 24: /SID52/gcgdata/geneseq/geneseq-n-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	ID	Description
1	22	68.8	28360	23 ABL06142	Drosophila melanog
2	21.4	66.9	1112	22 AAK13684	Enterococcus faeca
3	21.4	66.9	12426	22 AAS45482	Tumour suppressor
4	20.8	65.0	434	22 AAK57011	Human immune/haema
5	20.8	65.0	440	22 AAK69305	Human immune/haema
6	20.8	65.0	445	22 AAK66355	Human immune/haema
7	20.8	65.0	445	22 AAK66358	Human immune/haema
8	20.8	65.0	445	22 AAK66496	Human immune/haema
9	20.8	65.0	445	22 AAK66499	Human immune/haema

C 10	20.8	65.0	445	22	AAK69304	Human immune/haema
C 11	20.8	65.0	445	22	AAK69306	Human immune/haema
C 12	20.8	65.0	445	22	AAK69307	Human immune/haema
C 13	20.8	65.0	445	22	AAK84858	Human immune/haema
C 14	20.8	65.0	445	22	AAK84858	Tumour suppressor
C 15	20.8	65.0	13377	22	AAK46476	Human immune/haema
C 16	20.8	65.0	13377	24	ABL33463	Human immune syste
C 17	20.8	65.0	18501	23	ABL43029	Genomic sequence #
C 18	20.8	65.0	335913	22	AAK161371	Soybean 240017 reg
C 19	20.8	65.0	335913	22	AAK161372	Soybean 240017 reg
C 20	20.4	63.7	504	22	AAK283336	Genomic sequence #
C 21	20.4	63.7	504	22	AAK283336	Human immunoglobul
C 22	20.4	63.7	548	24	ABK53294	Human eosinophil-m
C 23	20.4	63.7	1972	22	AAK07834	Human secreted pro
C 24	20.4	63.7	2771	22	AAK07896	Human secreted pro
C 25	20.4	63.7	15141	23	ABL19378	Drosophila melanog
C 26	20.4	63.7	50000	24	ABL56202	Arabidopsis thalia
C 27	20	62.5	1140	21	AAK44850	AmEPV genome fragm
C 28	20	62.5	36445	23	ABL18116	Drosophila melanog
C 29	19.8	61.9	335	24	ABN18845	Human ORFX polynuc
C 30	19.8	61.9	471	22	AAS09349	Rac sequence diffe
C 31	19.8	61.9	562	24	ABK62832	Pneumocystis carin
C 32	19.8	61.9	621	12	AAK14598	Human colon cancer
C 33	19.8	61.9	760	21	AAA01825	P.carinii dihydrof
C 34	19.8	61.9	891	13	AAQ31427	Human polynucleoti
C 35	19.8	61.9	892	12	AAQ11961	Human ovarian anti
C 36	19.8	61.9	1936	22	AAI58011	Human secreted pro
C 37	19.8	61.9	2291	24	ABQ55062	Human polynucleoti
C 38	19.8	61.9	2776	20	AAK00678	Human secreted pro
C 39	19.8	61.9	2833	22	AAI59797	Human polynucleoti
C 40	19.8	61.9	3001	21	AAH51783	Chromosome 13q31-q
C 41	19.8	61.9	5945	24	ABL32084	Human immune syste
C 42	19.8	61.9	6337	23	ABL24528	Drosophila melanog
C 43	19.8	61.9	8477	23	ABL03628	Drosophila melanog
C 44	19.8	61.9	8801	22	AAK45436	Chemically pretrea
C 45	19.8	61.9	8801	24	ABL33740	Human immune syste

ALIGNMENTS

RESULT 1
ABL06142
ID ABL06142 standard; cDNA; 28360 BP.
XX ABL06142;
XX 26-MAR-2002 (first entry)
DT Drosophila melanogaster expressed polynucleotide SEQ ID NO 12908.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
PD 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP NY.
PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
DR P-PSDB; ABB62039.
XX New isolated nucleic acid detection reagent for detecting 1000 or more

diagnosing Enterococcal infection in an animal and monitoring progression of disease, and for identifying agents which can be used to modulate the growth or pathogenicity of *Enterococcus faecalis*, or other related organism, in vivo or in vitro. In particular the polypeptides encoded by the *Enterococcus faecalis* nucleotide sequences can be used in vaccines to prevent or attenuate an Enterococcal infection.

Sequence 1112 BP; 378 A; 171 C; 130 G; 428 T; 5 other;

Query Match 66.9%; Score 21.4; DB 20; Length 1112;
Best Local Similarity 80.6%; Pred. NO. 1.4e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATGTTGTTGTTT 31
||| ||||| ||||| ||| ||||| ||
Db 692 TAAACAAACATATATATATATGTTGTTT 662

RESULT 3

AAS46482/c
ID AAS46482 standard; DNA; 12426 BP.
XX
AAS46482:
AC AC
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #204.
DE
XX Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
XX Homo sapiens.
XX
XX WO200168912-A2.
PN
XX
XX 20-SEP-2001.
PD
XX
XX 15-MAR-2001; 2001WO-EP02955.
PF
XX
XX 15-MAR-2000; 2000DE-1013847.
PR
XX 06-APR-2000; 2000DE-1019058.
PR
XX 07-APR-2000; 2000DE-1019173.
PR
XX 30-JUN-2000; 2000DE-1032529.
PR
XX 01-SEP-2000; 2000DE-1043826.
PR
XX (EPIG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX WPI; 2001-602752/68.
DR
XX
XX Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -
XX
XX Claim 1; SEQ ID No 204; 27pp; English.
PS
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific

CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).

CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 12426 BP; 3791 A; 137 C; 2517 G; 5981 T; 0 other;

Query Match 66.98; Score 21.4; DB 22; Length 12426;
Best Local Similarity 80.68; Pred. No. 1.3e+02;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AATACACAAATATTATTGTGTGTTATTA 32

I | | | | | | | | | | | | | | | | | | |

Db 3282 ATTACACACAAATTAATTATATATATATA 3252

RESULT 4

AAK57011/c

ID AAK57011 standard; cDNA; 434 BP.

XX AC AAK57011;

XX DT 06-NOV-2001 (first entry)

XX DE Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:2071.

XX KW Human; Immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX OS Homo sapiens.

XX PN WO200157182-A2.

XX PD 09-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US01354.

XX PR 31-JAN-2000; 2000US-0179065.

XX PR 04-FEB-2000; 2000US-0180628.

XX PR 24-FEB-2000; 2000US-0184564.

XX PR 02-MAR-2000; 2000US-0186350.

XX PR 15-MAR-2000; 2000US-0189874.

XX PR 17-MAR-2000; 2000US-0190076.

XX PR 18-APR-2000; 2000US-0198123.

XX PR 19-MAY-2000; 2000US-0205515.

XX PR 07-JUN-2000; 2000US-0209467.

XX PR 28-JUN-2000; 2000US-0214886.

XX PR 30-JUN-2000; 2000US-0215135.

XX PR 07-JUL-2000; 2000US-0216647.

XX PR 07-JUL-2000; 2000US-0216880.

XX PR 11-JUL-2000; 2000US-0217487.

XX PR 14-JUL-2000; 2000US-0218290.

XX PR 26-JUL-2000; 2000US-0220963.

XX PR 26-JUL-2000; 2000US-0220964.

XX PR 14-AUG-2000; 2000US-0224518.

XX PR 14-AUG-2000; 2000US-0224519.

XX PR 14-AUG-2000; 2000US-0225213.

XX PR 14-AUG-2000; 2000US-0225214.

XX PR 14-AUG-2000; 2000US-0225266.

XX PR 14-AUG-2000; 2000US-0225267.

XX PR 14-AUG-2000; 2000US-0225268.

XX PR 14-AUG-2000; 2000US-0225270.

XX PR 14-AUG-2000; 2000US-0225271.

XX PR 14-AUG-2000; 2000US-0225447.

PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 12-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0233397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.


```
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 29-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 01-NOV-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-024617.
PR 01-NOV-2000; 2000US-0246174.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.

PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 03-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
XX Disclosure; SEQ ID NO 24117; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 440 BP; 140 A; 109 C; 94 G; 97 T; 0 other;

Query Match 65.0%; Score 20.8; DB 22; Length 440;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ATACAACACAAATATTAAATGTGTT 26
| | | | | | | | | | | | | | | | | |
Db 235 ATACAACACAAATGAATGTGTT 212

RESULT 6
AAK66355
ID AAK66355 standard; DNA; 445 BP.
XX
AC AAK66355;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human immune/hematopoietic antigen genomic sequence SEQ ID NO:21167.
XX
XX Human; immune; hematopoietic; immune/hematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX Homo sapiens.
OS
```

XX	PR	27-SEP-2000;	2000US-0235836.
PN	PR	29-SEP-2000;	2000US-0236327.
XX	PR	29-SEP-2000;	2000US-0236367.
PD	PR	29-SEP-2000;	2000US-0236368.
XX	PR	29-SEP-2000;	2000US-0236369.
PF	PR	29-SEP-2000;	2000US-0236370.
XX	PR	02-OCT-2000;	2000US-0236802.
XX	PR	02-OCT-2000;	2000US-0237037.
PR	PR	02-OCT-2000;	2000US-0237038.
PR	PR	02-OCT-2000;	2000US-0237039.
PR	PR	02-OCT-2000;	2000US-0237040.
PR	PR	13-OCT-2000;	2000US-0239935.
PR	PR	13-OCT-2000;	2000US-0239937.
PR	PR	20-OCT-2000;	2000US-0240960.
PR	PR	20-OCT-2000;	2000US-0241221.
PR	PR	20-OCT-2000;	2000US-0241785.
PR	PR	20-OCT-2000;	2000US-0241786.
PR	PR	20-OCT-2000;	2000US-0241787.
PR	PR	20-OCT-2000;	2000US-0241808.
PR	PR	20-OCT-2000;	2000US-0241809.
PR	PR	20-OCT-2000;	2000US-0241826.
PR	PR	01-NOV-2000;	2000US-0244617.
PR	PR	08-NOV-2000;	2000US-0246474.
PR	PR	08-NOV-2000;	2000US-0246475.
PR	PR	08-NOV-2000;	2000US-0246476.
PR	PR	08-NOV-2000;	2000US-0246477.
PR	PR	08-NOV-2000;	2000US-0246478.
PR	PR	08-NOV-2000;	2000US-0246523.
PR	PR	08-NOV-2000;	2000US-0246524.
PR	PR	08-NOV-2000;	2000US-0246525.
PR	PR	08-NOV-2000;	2000US-0246526.
PR	PR	08-NOV-2000;	2000US-0246527.
PR	PR	08-NOV-2000;	2000US-0246528.
PR	PR	08-NOV-2000;	2000US-0246532.
PR	PR	08-NOV-2000;	2000US-0246609.
PR	PR	08-NOV-2000;	2000US-0246610.
PR	PR	08-NOV-2000;	2000US-0246611.
PR	PR	08-NOV-2000;	2000US-0246613.
PR	PR	17-NOV-2000;	2000US-0249207.
PR	PR	17-NOV-2000;	2000US-0249208.
PR	PR	17-NOV-2000;	2000US-0249209.
PR	PR	17-NOV-2000;	2000US-0249210.
PR	PR	17-NOV-2000;	2000US-0249211.
PR	PR	17-NOV-2000;	2000US-0249212.
PR	PR	17-NOV-2000;	2000US-0249213.
PR	PR	17-NOV-2000;	2000US-0249214.
PR	PR	17-NOV-2000;	2000US-0249215.
PR	PR	17-NOV-2000;	2000US-0249216.
PR	PR	17-NOV-2000;	2000US-0249217.
PR	PR	17-NOV-2000;	2000US-0249218.
PR	PR	17-NOV-2000;	2000US-0249244.
PR	PR	17-NOV-2000;	2000US-0249245.
PR	PR	17-NOV-2000;	2000US-0249264.
PR	PR	17-NOV-2000;	2000US-0249265.
PR	PR	17-NOV-2000;	2000US-0249297.
PR	PR	17-NOV-2000;	2000US-0249299.
PR	PR	17-NOV-2000;	2000US-0249300.
PR	PR	01-DEC-2000;	2000US-0250160.
PR	PR	01-DEC-2000;	2000US-0250391.
PR	PR	05-DEC-2000;	2000US-0251030.
PR	PR	05-DEC-2000;	2000US-0251988.
PR	PR	05-DEC-2000;	2000US-0256719.
PR	PR	06-DEC-2000;	2000US-0251479.
PR	PR	08-DEC-2000;	2000US-0251856.
PR	PR	08-DEC-2000;	2000US-0251868.
PR	PR	08-DEC-2000;	2000US-0251869.
PR	PR	08-DEC-2000;	2000US-0251989.
PR	PR	08-DEC-2000;	2000US-0251990.
PR	PR	11-DEC-2000;	2000US-0254097.
PR	PR	05-JAN-2001;	2001US-0259678.
XX	XX		
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX	XX		


```
PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis.
XX Disclosure; SEQ ID NO 21167; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX Sequence 445 BP; 98 A; 94 C; 112 G; 141 T; 0 other;
SQ
Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ATACACACAAATATTATTGTTGTT 26
|||||
Db 211 ATACACACAAATGAATGTTGTT 234
|||||
RESULT 7
AAK6358
ID AAK66358 standard; DNA; 445 BP.
XX
AC AAK66358;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human Immune/hematopoietic antigen genomic sequence SEQ ID NO:21170.
XX
KW Human; immune; haematopoietic; immune/hematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 13-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
```

20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 03-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX Disclosure; SEQ ID NO 21170; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
CC amino acid sequences given in AM82170 to AM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting

CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/hematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 445 BP; 98 A; 94 C; 112 G; 141 T; 0 other;
Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 3 ATACAACACACATATTAATTGTGTT 26
Db 211 ATACAACACACAAATGATGTGTT 234
RESULT 8
AAK66496
ID AAK66496 standard; DNA; 445 BP.
XX AC
XX AAK66496;
XX 06-NOV-2001 (first entry)
XX Human immune/hematopoietic antigen genomic sequence SEQ ID NO:21308.
DE Human; immune; hematopoietic; immune/hematopoietic antigen; cancer;
KW cytotstatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
XX WO200157182-A2.
XX 09-AUG-2001.
XX 17-JAN-2001; 2001WO-US01354.
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.

```
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234597.
PR 25-SEP-2000; 2000US-0234598.
PR 25-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 06-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251866.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI WPI; 2001-483426/52.
XX
DR Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 21308; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent
XX diagnose and treat immune/hematopoietic-related diseases, especially
XX cancers and cancer metastases of hematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/hematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 445 BP; 98 A; 94 C; 112 G; 141 T; 0 other;
Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ATACACACAATATTAATGTGTT 26
||||||| || |||||
Db 211 ATACACACAATGAATGTGTT 234
RESULT 9
```

AAK66499
ID AAK66499 standard; DNA; 445 BP.
AC AAK66499;
XX
XX
DT 06-NOV-2001 (first entry)
XX
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:21311.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
KW
XX
XX Homo sapiens.
XX
PN WO200157182-A2.
PN
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 28-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225277.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234397.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.

```
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX
XX Disclosure; SEQ ID NO 21311; 3071pp + Sequence Listing; English.
XX
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 445 BP; 98 A; 94 C; 112 G; 141 T; 0 other;
XX
XX Query Match 65.0%; Score 20.8; DB 22; Length 445;
XX Best Local Similarity 91.7%; Pred. No. 2.2e+02;
XX Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 3 ATACACACAAATATTATTGTT 26
XX ||||||||| |||||||||
XX Db 211 ATACACACAAATGATTGTT 234
XX
XX RESULT 10
XX AAK69304/c
XX ID AAK69304 standard; DNA; 445 BP.
XX
XX AC AAK69304;
XX
XX XX
XX DT 06-NOV-2001 (first entry)
XX
XX XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24116.
XX DE
XX DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX KW
XX OS Homo sapiens.
XX
XX XX WO200157182-A2.
XX PN
XX XX
XX PD 09-AUG-2001.
XX
XX XX 17-JAN-2001; 2001WO-US01354.
XX
XX XX
```

```
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
```

```
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249246.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PR
PR (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI; 2001-483426/52.
XX
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
```

```
XX Disclosure; SEQ ID NO 24116; 3071pp + Sequence Listing; English.
PS
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
XX Sequence 445 BP; 141 A; 112 C; 94 G; 98 T; 0 other;
SQ
Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ATACACACACATATTATTGTTGTT 26
Db 235 ATACACACACAAATGAATTGTTGTT 212
RESULT 11
AAK69306/c
ID AAK69306 standard; DNA; 445 BP.
XX
AC AAK69306;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24118.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
PN WO200157182-A2.
XX
PD 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 02-MAR-2000; 2000US-0184664.
PR 16-MAR-2000; 2000US-0186350.
PR 17-MAR-2000; 2000US-0189874.
PR 18-APR-2000; 2000US-0190076.
PR 19-MAY-2000; 2000US-0198123.
PR 07-JUN-2000; 2000US-0205515.
PR 28-JUN-2000; 2000US-0209467.
PR 30-JUN-2000; 2000US-0214886.
PR 07-JUL-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 14-JUL-2000; 2000US-0217496.
PR 26-JUL-2000; 2000US-0218290.
PR 14-AUG-2000; 2000US-0220963.
PR 14-AUG-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
```

PR	14-AUG-2000;	2000US-022521310;
PR	14-AUG-2000;	2000US-022521414;
PR	14-AUG-2000;	2000US-022525666;
PR	14-AUG-2000;	2000US-022526677;
PR	14-AUG-2000;	2000US-022526688;
PR	14-AUG-2000;	2000US-022526710;
PR	14-AUG-2000;	2000US-022525727;
PR	14-AUG-2000;	2000US-022535474;
PR	14-AUG-2000;	2000US-022537577;
PR	14-AUG-2000;	2000US-022537588;
PR	14-AUG-2000;	2000US-022537589;
PR	18-AUG-2000;	2000US-022562729;
PR	22-AUG-2000;	2000US-02266681;
PR	22-AUG-2000;	2000US-022668688;
PR	22-AUG-2000;	2000US-022718281;
PR	23-AUG-2000;	2000US-022700909;
PR	30-SEP-2000;	2000US-022892424;
PR	01-SEP-2000;	2000US-0229287;
PR	01-SEP-2000;	2000US-0229343;
PR	01-SEP-2000;	2000US-0229344;
PR	01-SEP-2000;	2000US-0229345;
PR	05-SEP-2000;	2000US-0229350909;
PR	05-SEP-2000;	2000US-02295113;
PR	06-SEP-2000;	2000US-0230437;
PR	08-SEP-2000;	2000US-0230438;
PR	08-SEP-2000;	2000US-0231242;
PR	08-SEP-2000;	2000US-0231243;
PR	08-SEP-2000;	2000US-0231244;
PR	08-SEP-2000;	2000US-0231413;
PR	08-SEP-2000;	2000US-0231414;
PR	08-SEP-2000;	2000US-0232080;
PR	12-SEP-2000;	2000US-0232081;
PR	14-SEP-2000;	2000US-0231968;
PR	14-SEP-2000;	2000US-0232397;
PR	14-SEP-2000;	2000US-0232398;
PR	14-SEP-2000;	2000US-0232399;
PR	14-SEP-2000;	2000US-0232400;
PR	14-SEP-2000;	2000US-0232401;
PR	14-SEP-2000;	2000US-0233063;
PR	14-SEP-2000;	2000US-0233064;
PR	21-SEP-2000;	2000US-0233223;
PR	21-SEP-2000;	2000US-0234274;
PR	23-SEP-2000;	2000US-0234977;
PR	23-SEP-2000;	2000US-0234998;
PR	26-SEP-2000;	2000US-0235484;
PR	27-SEP-2000;	2000US-0235834;
PR	27-SEP-2000;	2000US-0235836;
PR	29-SEP-2000;	2000US-0236327;
PR	29-SEP-2000;	2000US-0236367;
PR	29-SEP-2000;	2000US-0236368;
PR	29-SEP-2000;	2000US-0236369;
PR	29-SEP-2000;	2000US-0236370;
PR	02-OCT-2000;	2000US-0236802;
PR	02-OCT-2000;	2000US-0237037;
PR	02-OCT-2000;	2000US-0237038;
PR	02-OCT-2000;	2000US-0237039;
PR	02-OCT-2000;	2000US-0237040;
PR	13-OCT-2000;	2000US-0239935;
PR	13-OCT-2000;	2000US-0239937;
PR	20-OCT-2000;	2000US-0240960;
PR	20-OCT-2000;	2000US-0241222;
PR	20-OCT-2000;	2000US-0241785;
PR	20-OCT-2000;	2000US-0241786;
PR	20-OCT-2000;	2000US-0241787;
PR	20-OCT-2000;	2000US-0241808;
PR	20-OCT-2000;	2000US-0241809;
PR	20-OCT-2000;	2000US-0241826;
PR	01-NOV-2000;	2000US-0244617;
PR	08-NOV-2000;	2000US-0246474;
PR	08-NOV-2000;	2000US-0246475;
PR	08-NOV-2000;	2000US-0246476;
PR	08-NOV-2000;	2000US-0246477;
PR	08-NOV-2000;	2000US-0246478;
PR	08-NOV-2000;	2000US-0246479;

```
XX SQ Sequence 445 BP; 141 A; 112 C; 94 G; 98 T; 0 other;
Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ATACACACAATATTATTGTTT 26
||||| ||| ||| ||| |||
Db 235 ATACACACAAAATGAATTGTTT 212

RESULT 12
AAK69307/c
ID AAK69307 standard; DNA; 445 BP.
XX AC AAK69307;
XX DT 06-NOV-2001 (first entry)
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:24119.
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX OS Homo sapiens.
XX PN WO200157182-A2.
XX PD 09-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US01354.
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
XX PR 19-MAY-2000; 2000US-0205515.
XX PR 07-JUN-2000; 2000US-0209467.
XX PR 28-JUN-2000; 2000US-0214886.
XX PR 30-JUN-2000; 2000US-0215135.
XX PR 07-JUL-2000; 2000US-0216647.
XX PR 07-JUL-2000; 2000US-0216880.
XX PR 11-JUL-2000; 2000US-0217487.
XX PR 11-JUL-2000; 2000US-0217496.
XX PR 14-JUL-2000; 2000US-0218290.
XX PR 26-JUL-2000; 2000US-0220963.
XX PR 26-JUL-2000; 2000US-0220964.
XX PR 14-AUG-2000; 2000US-0224518.
XX PR 14-AUG-2000; 2000US-0224519.
XX PR 14-AUG-2000; 2000US-0225213.
XX PR 14-AUG-2000; 2000US-0225214.
XX PR 14-AUG-2000; 2000US-0225266.
XX PR 14-AUG-2000; 2000US-0225267.
XX PR 14-AUG-2000; 2000US-0225268.
XX PR 14-AUG-2000; 2000US-0225270.
XX PR 14-AUG-2000; 2000US-0225447.
XX PR 14-AUG-2000; 2000US-0225757.
XX PR 14-AUG-2000; 2000US-0225758.
XX PR 14-AUG-2000; 2000US-0225759.
XX PR 18-AUG-2000; 2000US-0226279.
XX PR 22-AUG-2000; 2000US-0226681.
XX PR 22-AUG-2000; 2000US-0226688.
XX PR 22-AUG-2000; 2000US-0227182.
XX PR 23-AUG-2000; 2000US-0227009.
XX PR 30-AUG-2000; 2000US-0228924.
XX PR 01-SEP-2000; 2000US-0229287.
XX PR 01-SEP-2000; 2000US-0229343.
XX PR 01-SEP-2000; 2000US-0229344.
XX PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244517.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
```



```

PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249219.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and
XX metastasis -
XX Disclosure: SEQ ID NO 39670; 3071pp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patient's own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX the nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAK82169
XX represent sequences used in the exemplification of the present invention.
XX
XX SQ Sequence 445 BP; 98 A; 94 C; 112 G; 141 T; 0 other;
Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e+02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 ATACACACAAATATTAAATGTGTT 26
Db 211 ATACACACAAATGAATGTGTT 234
RESULT 14
AAK84862
ID AAK84862 standard; DNA; 445 BP.
XX
XX AC AAK84862;
XX
XX DT 07-NOV-2001 (first entry)
XX
XX DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:39674.
XX
XX KW Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
XX cytostatic; gene therapy; vaccine; metastasis; ds.
XX
XX OS Homo sapiens.
XX
XX PN WO200157182-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US01354.
XX
XX PR 31-JAN-2000; 2000US-0179065.
XX PR 04-FEB-2000; 2000US-0180628.
XX PR 24-FEB-2000; 2000US-0184664.
XX PR 02-MAR-2000; 2000US-0186350.
XX PR 16-MAR-2000; 2000US-0189874.
XX PR 17-MAR-2000; 2000US-0190076.
XX PR 18-APR-2000; 2000US-0198123.
```

PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0215647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226581.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231245.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 17-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249266.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
(HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX
PI
PI
XX
XX
DR
XX
XX
PT
PT
PT
XX
XX
PS
XX
XX
CC
CC
CC
CC
Nucleic acids encoding human immune/hematopoietic antigen polypeptides, useful for preventing, diagnosing and/or treating cancers and metastasis -
Disclosure: SEQ ID NO 39674; 3071pp + Sequence Listing; English.
AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I) amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic activity, and can be used in gene therapy and vaccine production. (I) proteins and polynucleotides may be used in the prevention, diagnosis and

CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAK82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 445 BP; 98 A; 94 C; 112 G; 141 T; 0 other;

Query Match 65.0%; Score 20.8; DB 22; Length 445;
Best Local Similarity 91.7%; Pred. No. 2.2e-02;
Matches 22; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 ATACAACACAATATTAAATGTGTT 26
|||||
Db 211 ATACAACACAATGAATGTGTT 234

RESULT 15

AA546476
ID AAS46476 standard; DNA; 13377 BP.
XX
AC AAS46476;

XX
DT 18-DEC-2001 (first entry)

XX
DE Tumour suppressor gene derived chemically modified sequence #198.

XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.

XX
OS Homo sapiens.

XX
PN WO200168912-A2.

XX
PD 20-SEP-2001.

XX
PF 15-MAR-2001; 2001WO-EP02955.

XX
PR 15-MAR-2000; 2000DE-1013847.

XX
PR 06-APR-2000; 2000DE-1019058.

XX
PR 07-APR-2000; 2000DE-1019173.

XX
PR 30-JUN-2000; 2000DE-1032529.

XX
PR 01-SEP-2000; 2000DE-1043826.

XX
PA (EPIC-) EPIGENOMICS AG.

XX
PI Olek A, Piepenbrock C, Berlin K;

XX
WPI; 2001-602752/68.

XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer -

XX
PS Claim 1; SEQ ID NO 198; 27pp; English.

XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may

CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 13377 BP; 3714 A; 146 C; 2702 G; 6815 T; 0 other;

Query Match 65.0%; Score 20.8; DB 22; Length 13377;
Best Local Similarity 78.1%; Pred. No. 2.1e-02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAATACACACAATATTAAATGTGTGTTATTA 32
|||||
Db 5498 TTATATAAAGAATATTAAATGCGTGATTA 5529

Search completed: February 13, 2003, 01:45:42
Job time : 300 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2003, 01:38:51 ; Search time 66 Seconds
(without alignments)
148.692 Million cell updates/sec

Title: US-09-895-435-3
Perfect score: 32
Sequence: 1 taatacacacaataattgtgtgtatta 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19.8	61.9	891	6	Patent No. 5164490-2
2	19.2	60.0	36	3	US-08-770-941-13
3	19.2	60.0	342	4	US-09-134-001C-2702
4	19.2	60.0	360	3	US-08-935-263-17
5	19.2	60.0	360	4	US-09-594-185-17
6	18.8	58.8	266	2	US-08-687-080-86
7	18.8	58.8	564	4	US-08-984-320-1
8	18.8	58.8	564	4	US-08-487-087A-1
9	18.8	58.8	564	4	US-08-397-653B-1
10	18.8	58.8	564	6	5175095-1
11	18.8	58.8	2132	3	US-08-844-188-39
12	18.8	58.8	2132	3	US-09-378-088A-39
13	18.8	58.8	3528	4	US-08-984-320-2
14	18.8	58.8	3528	4	US-08-487-087A-2
15	18.8	58.8	4383	4	US-08-397-653B-2
16	18.8	58.8	4383	6	5175095-4
17	18.8	58.8	4383	6	5177307-1
18	18.8	58.8	4833	4	US-09-066-047-1
19	18.6	58.1	3155	2	US-08-591-629-7
20	18.6	58.1	168575	4	US-09-426-290-1
21	18.4	57.5	3561	1	US-08-485-568A-3
22	18.4	57.5	3561	1	US-08-357-698-5
23	18.4	57.5	3561	2	US-08-590-554A-3
24	18.4	57.5	3561	2	US-09-184-223-3
25	18.4	57.5	3561	5	PCT-US93-12682-5
26	18.4	57.5	4435	1	US-08-308-872B-5
27	18.2	56.9	1107	4	US-09-173-300-50

28	18.2	56.9	1483	1	US-07-966-187-11	Sequence 11, Appl
29	18.2	56.9	1483	1	US-07-966-187-13	Sequence 13, Appl
30	18.2	56.9	2301	4	US-09-561-825-1	Sequence 1, Appl
31	18.2	56.9	2351	4	US-09-561-825-26	Sequence 26, Appl
32	18.2	56.9	2351	4	US-09-561-825-29	Sequence 29, Appl
33	18.2	56.9	2351	4	US-09-561-825-27	Sequence 27, Appl
34	18.2	56.9	2351	4	US-09-561-825-28	Sequence 28, Appl
35	18.2	56.9	6343	3	US-08-581-148C-30	Sequence 30, Appl
36	18.2	56.9	11721	4	US-09-026-039-3	Sequence 3, Appl
37	18.2	56.9	11721	4	US-09-026-039-3	Sequence 3, Appl
38	18	56.2	652	4	US-08-936-165A-105	Sequence 105, App
39	18	56.2	712	1	US-08-410-804-4	Sequence 4, Appl
40	18	56.2	712	1	US-08-259-514-4	Sequence 4, Appl
41	18	56.2	712	2	US-08-858-311-4	Sequence 4, Appl
42	18	56.2	747	1	US-08-680-726A-87	Sequence 87, Appl
43	18	56.2	747	3	US-09-092-408-87	Sequence 87, Appl
44	18	56.2	749	1	US-08-680-726A-85	Sequence 85, Appl
45	18	56.2	749	1	US-08-680-726A-86	Sequence 86, Appl

ALIGNMENTS

RESULT 1
5164490-2
; Patent No. 5164490
; APPLICANT: SANTI, DANIEL V.; EDMAN, JEFFREY; EDMAN, URSULA
; TITLE OF INVENTION: PNEUMOCYSTIS CARINII DIHYDROFOLATE
; REDUCTASE GENE AND METHODS FOR ITS USE
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/437,511
; FILING DATE: 15-NOV-1989
; SEQ ID NO:2:
; LENGTH: 891
5164490-2

Query Match 61.9%; Score 19.8; DB 6; Length 891;
Best Local Similarity 77.4%; Pred. No. 48;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAATACACACAATATTATTTGTTGTATT 31
||||| - ||||| ||||| |||||
Db 507 TAATATACAGGATTCATTTGTGATGATT 537

RESULT 2
US-08-770-941-13
; Sequence 13, Application US/08770941
; Patent No. 6025133
; GENERAL INFORMATION:
; APPLICANT: Stull, Paul D.
; APPLICANT: Myers, Kristi K.
; APPLICANT: Becker, Michael M.
; TITLE OF INVENTION: TARGET-TRIGGERED
; AMPLIFICATION
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,941

```

; CURRENT APPLICATION NUMBER: US/08/935,263A
; CURRENT FILING DATE: 1997-09-22
; EARLIER APPLICATION NUMBER: EP 96115540.5
; EARLIER FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Kurthia sp.
; FEATURE:
; NAME/KEY: RBS
; LOCATION: Complement((67)..(76))
; FEATURE:
; NAME/KEY: -35_signal
; LOCATION: (210)..(215)
; FEATURE:
; NAME/KEY: -10_signal
; LOCATION: (234)..(239)
; FEATURE:
; NAME/KEY: -10_signal
; LOCATION: Complement((235)..(240))
; FEATURE:
; NAME/KEY: RBS
; LOCATION: (289)..(293)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (302)..(358)
; OTHER INFORMATION: Partial sequence of ORF2.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (125)..(164)
; OTHER INFORMATION: BOX1 - inverted repeat
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (244)..(283)
; OTHER INFORMATION: BOX2 - inverted repeat
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: Complement((1)..(58))
; OTHER INFORMATION: Partial sequence of ORF1.
;
US-08-935-263-17

Query Match          60.0%; Score 19.2; DB 3; Length 360;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 1 TAAATACACACAATATTAAATGTGTGTGTTATTA 32
    |||| |||| || |||| |||| |||| ||||
Db 207 TAAATGAACAATATTTTATTTGTGTGTTATTA 238

RESULT 5
US-09-594-185-17
; Sequence 17, Application US/09594185
; Patent No. 6365388
; GENERAL INFORMATION:
; APPLICANT: Furuichi, Yasuhiro
; APPLICANT: Hoshino, Tatsuo
; APPLICANT: Kimura, Hitoshi
; APPLICANT: Kiyasu, Tatsuya
; APPLICANT: Nagahashi, Yoshie
; TITLE OF INVENTION: BIOTIN BIOSYNTHETIC GENES
; FILE REFERENCE: Biotin Genes
; CURRENT APPLICATION NUMBER: US/09/594,185
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 08/935,263
; PRIOR FILING DATE: 1997-09-22
; PRIOR APPLICATION NUMBER: EP 96115540.5
; PRIOR FILING DATE: 1996-09-27
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17

```

```
;
; LENGTH: 360
; TYPE: DNA
; ORGANISM: Kurthia sp.
; FEATURE:
; NAME/KEY: RBS
; LOCATION: Complement((67)..(76))
; NAME/KEY: -35_signal
; LOCATION: (210)..(215)
; NAME/KEY: -10_signal
; LOCATION: (234)..(239)
; NAME/KEY: -10_signal
; LOCATION: Complement((235)..(240))
; NAME/KEY: RBS
; LOCATION: (289)..(293)
; NAME/KEY: misc_feature
; LOCATION: (302)..(358)
; OTHER INFORMATION: Partial sequence of ORF2.
; NAME/KEY: misc_feature
; LOCATION: (125)..(164)
; OTHER INFORMATION: BOX1 - inverted repeat
; NAME/KEY: misc_feature
; LOCATION: (244)..(283)
; OTHER INFORMATION: BOX2 - inverted repeat
; NAME/KEY: misc_feature
; LOCATION: Complement((1)..(58))
; OTHER INFORMATION: Partial sequence of ORF1.
US-09-594-185-17
```

```
Query Match 60.0%; Score 19.2; DB 4; Length 360;
Best Local Similarity 75.0%; Pred. No. 80;
Matches 24; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 TAATACACACATATTAATTGTTGTATTATTA 32
||||| ||||| ||||| ||||| |||||
Db 207 TAATGAACAATATTTATTTGTTGTATTATTA 238
```

```
RESULT 6
US-08-687-080-86
; Sequence 86, Application US/08687080
; Patent No. 5965427
; GENERAL INFORMATION:
; APPLICANT: Gregory Dolganov
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof
; NUMBER OF SEQUENCES: 175
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: 350 Cambridge Avenue, Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,080
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/592,126
; FILING DATE: 26-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sholtz, Charles K.
; REGISTRATION NUMBER: 38,615
; REFERENCE/DOCKET NUMBER: 4600-0111.30
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 86:
; SEQUENCE CHARACTERISTICS:
```

```
;
; LENGTH: 266 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 3' END OF INTRON 13 OF RAD50 GENOMIC
; INDIVIDUAL ISOLATE: SEQUENCE
US-08-687-080-86

Query Match 58.8%; Score 18.8; DB 2; Length 266;
Best Local Similarity 76.7%; Pred. No. 11e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACACACATATTAATTGTTGTATT 31
||| ||||| ||||| ||||| |||||
Db 160 AAAAGAACACAAATGTCACCTTCTGTGTATT 189

RESULT 7
US-08-984-320-1/c
; Sequence 1, Application US/08984320
; Patent No. 6222097
; GENERAL INFORMATION:
; APPLICANT: McBride, Kevin E.
; APPLICANT: Stalker, David M.
; TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,087
; FILING DATE: 07-JUN-95
; APPLICATION NUMBER: USSN 07/998,158
; FILING DATE: 29-DEC-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/554,195
; FILING DATE: 17-JUL-90
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/382,518
; FILING DATE: 19-JUL-89
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 91-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
```

Thu Feb 13 12:34:56 2003

us-09-895-435-3.rni

```
US-08-984-320-1
Query Match          58.8%; Score 18.8; DB 4; Length 564;
Best Local Similarity 76.7%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACAAACAATATTAATTGTTGTATT 31
   || ||||| ||||| ||||| ||||| ||
Db 423 AAGGCAACACTTATTAATTGTTGTTTT 394

RESULT 8
US-08-487-087A-1/c
; Sequence 1, Application US/08487087A
; Patent No. 6268546
; GENERAL INFORMATION:
; APPLICANT: McBride, Kevin E.
; APPLICANT: Stalker, David M.
; TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
; COMPUTER: Apple Macintosh 7.1
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,087A
; FILING DATE: 07-JUN-95
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USN 07/998,158
; FILING DATE: 29-DEC-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/554,195
; FILING DATE: 17-JUL-90
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/382,518
; FILING DATE: 19-JUL-89
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 91-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-1510
; TELEFAX: (916) 753-6313
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
US-08-487-087A-1
Query Match          58.8%; Score 18.8; DB 4; Length 564;
Best Local Similarity 76.7%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACAAACAATATTAATTGTTGTATT 31
   || ||||| ||||| ||||| ||||| ||
Db 423 AAGGCAACACTTATTAATTGTTGTTTT 394

RESULT 9
US-08-397-653B-1/c
; Sequence 1, Application US/08397653B
; Patent No. 6329570
; GENERAL INFORMATION:
; APPLICANT: Martineau, Belinda
; TITLE OF INVENTION: COTTON MODIFICATION USING
; TITLE OF INVENTION: OVARY-TISSUE TRANSCRIPTIONAL
; TITLE OF INVENTION: FACTORS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.0 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.0
; SOFTWARE: Microsoft Word 5.1a
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/397,653B
; FILING DATE: 28-FEB-1995
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Elizabeth Lassen
; REGISTRATION NUMBER: 31,845
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
US-08-397-653B-1
Query Match          58.8%; Score 18.8; DB 4; Length 564;
Best Local Similarity 76.7%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACAAACAATATTAATTGTTGTATT 31
   || ||||| ||||| ||||| ||||| ||
Db 423 AAGGCAACACTTATTAATTGTTGTTTT 394

RESULT 10
5175095-1/c
; Patent No. 5175095
; APPLICANT: Martineau, Belinda M.; Houck, Catherine M.
; TITLE OF INVENTION: OVARY TISSUE TRANSCRIPTIONAL FACTORS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/554,195
; FILING DATE: 17-JUL-1990
; SEQ ID NO: 1:
; LENGTH: 564
5175095-1
Query Match          58.8%; Score 18.8; DB 6; Length 564;
Best Local Similarity 76.7%; Pred. No. 1.1e+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACAAACAATATTAATTGTTGTATT 31
   || ||||| ||||| ||||| ||||| ||
Db 423 AAGGCAACACTTATTAATTGTTGTTTT 394
```


Db 423 AAGCAACACTTATTAATTTGTTGTTT 394

RESULT 11

US-08-844-188-39/C
; Sequence 39, Application US/08844188
; Patent No. 6127180
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy
; APPLICANT: Schwab, George E.
; TITLE OF INVENTION: Pesticidal Toxins
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik, Lloyd & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: USA
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,188
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/633,993
; FILING DATE: 19-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanders, Jay M.
; REGISTRATION NUMBER: 39,355
; REFERENCE/DOCKET NUMBER: MA-703C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 39:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2132 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-844-188-39

Query Match 58.8%; Score 18.8; DB 3; Length 2132;
Best Local Similarity 76.7%; Pred. No. 1.le+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACACACAATTAATTTGTTGTTATT 31
||||| | ||||| ||| | |||||
Db 724 AATACCAAAAATATTTATTTTGGTATT 695

RESULT 12

US-09-378-088A-39/C
; Sequence 39, Application US/09378088A
; Patent No. 6372480
; GENERAL INFORMATION:
; APPLICANT: Narva, Kenneth E.
; APPLICANT: Schnepf, H. Ernest
; APPLICANT: Knuth, Mark
; APPLICANT: Pollard, Michael R.
; APPLICANT: Cardineau, Guy A.
; APPLICANT: Schwab, George E.
; APPLICANT: Michaels, Tracy E.

; APPLICANT: Finstad Lee, Stacy
; APPLICANT: Burmeister, Paula
; APPLICANT: Dojillo, Joanna
; TITLE OF INVENTION: Pesticidal Proteins
; FILE REFERENCE: MA703C2
; CURRENT APPLICATION NUMBER: US/09/378,088A
; CURRENT FILING DATE: 1999-08-20
; PRIOR APPLICATION NUMBER: US 08/633,993
; PRIOR FILING DATE: 1996-04-19
; PRIOR APPLICATION NUMBER: US 08/844,188
; PRIOR FILING DATE: 1997-04-18
; NUMBER OF SEQ ID NOS: 130
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Bacillus thuringiensis
US-09-378-088A-39

Query Match 58.8%; Score 18.8; DB 4; Length 2132;
Best Local Similarity 76.7%; Pred. No. 1.le+02;
Matches 23; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 2 AATACACACAATTAATTTGTTGTTATT 31
||||| | ||||| ||| | |||||
Db 724 AATACCAAAAATATTTATTTTGGTATT 695

RESULT 13

US-08-984-320-2/C
; Sequence 2, Application US/08984320
; Patent No. 6222097
; GENERAL INFORMATION:
; APPLICANT: McBride, Kevin E.
; APPLICANT: Stalker, David M.
; TITLE OF INVENTION: Use of Ovary-Tissue Transcriptional Factors
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Calgene, Inc.
; STREET: 1920 Fifth Street
; CITY: Davis
; STATE: CA
; COUNTRY: USA
; ZIP: 95616
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4 MB
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.1
; SOFTWARE: Microsoft Word 5.1 (a)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,320
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,087
; FILING DATE: 07-JUN-95
; APPLICATION NUMBER: USSN 07/998,158
; FILING DATE: 29-DEC-92
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/554,195
; FILING DATE: 17-JUL-90
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 07/382,518
; FILING DATE: 19-JUL-89
; ATTORNEY/AGENT INFORMATION:
; NAME: Donna E. Scherer
; REGISTRATION NUMBER: 34,719
; NAME: Carl J. Schwedler
; REGISTRATION NUMBER: 36,924
; REFERENCE/DOCKET NUMBER: CGNE 91-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (916) 753-6313
; TELEFAX: (916) 753-1510

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2003, 01:41:21 : Search time 37 seconds
(without alignments)
440.496 Million cell updates/sec

Title: US-09-895-435-3

Perfect score: 32

Sequence: I taatacacacaataattgtgtgtatta 32

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 424239 seqs, 254661826 residues

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
1: /cgn2_5/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	21.4	66.9	1112	10	US-09-070-927A-747
2	21.4	66.9	2000	9	US-09-938-842A-4932
3	20.8	65.0	1876	9	US-09-938-842A-4407
C 4	20.8	65.0	18501	10	US-09-764-847-1916
5	20.8	65.0	335913	9	US-09-754-853A-2
6	20.8	65.0	335913	9	US-09-754-853A-3
7	20.4	63.7	504	10	US-09-764-860-770
C 8	20.2	63.1	448	10	US-09-960-352-11904
9	19.8	61.9	562	10	US-09-917-800A-739
C 10	19.8	61.9	1936	9	US-10-098-841-217
11	19.8	61.9	2776	9	US-09-984-245-78
12	19.8	61.9	127197	9	US-09-754-853A-1
C 13	19.8	61.9	148567	9	US-10-254-869-3
C 14	19.8	61.9	148567	10	US-09-801-876B-3
15	19.6	61.3	357	10	US-09-770-791-596
16	19.6	61.3	502	10	US-09-764-853-15
17	19.6	61.3	513509	9	US-09-754-853A-4
C 18	19.4	60.6	267	10	US-09-960-352-3174
19	19.4	60.6	180216	10	US-09-835-232-6

20	19.2	60.0	360	12	US-10-033-078-17	Sequence 17, Appl
C 21	19.2	60.0	374	9	US-09-954-531-559	Sequence 559, App
C 22	19.2	60.0	26225	10	US-09-764-889-1276	Sequence 1276, Ap
C 23	19.2	59.4	4565	9	US-10-138-221-8	Sequence 8, Appl
C 24	19.2	59.4	4565	10	US-09-794-384A-3	Sequence 3, Appl
C 25	19.2	59.4	4810	9	US-10-138-221-6	Sequence 6, Appl
26	19.2	59.4	25580	10	US-09-070-927A-19	Sequence 19, Appl
27	19.2	59.4	302250	10	US-09-962-832-154	Sequence 154, App
C 28	18.8	58.8	334	10	US-09-867-701-674	Sequence 674, App
C 29	18.8	58.8	564	10	US-09-777-347-1	Sequence 1, Appl
30	18.8	58.8	669	9	US-09-764-904-137	Sequence 137, App
31	18.8	58.8	669	10	US-09-764-860-1198	Sequence 1198, Ap
C 32	18.8	58.8	2000	9	US-09-938-842A-4906	Sequence 4906, Ap
C 33	18.8	58.8	3528	10	US-09-777-347-2	Sequence 2, Appl
C 34	18.8	58.8	74586	10	US-09-781-558-3	Sequence 3, Appl
C 35	18.8	58.8	74586	10	US-09-781-558-3	Sequence 4375, Ap
C 36	18.6	58.1	440	10	US-09-864-761-4375	Sequence 3329, A
37	18.6	58.1	591	10	US-09-864-761-33229	Sequence 3970, Ap
C 38	18.6	58.1	2000	9	US-09-938-842A-3970	Sequence 102, App
39	18.6	58.1	3029	9	US-09-974-298-102	Sequence 196, App
40	18.4	57.5	2004	10	US-09-887-576-196	Sequence 222, App
C 41	18.2	56.9	396	10	US-09-960-352-222	Sequence 2087, App
C 42	18.2	56.9	418	10	US-09-960-352-2087	Sequence 138, App
C 43	18.2	56.9	462	10	US-09-880-107-138	Sequence 10268, A
C 44	18.2	56.9	481	10	US-09-864-761-10268	Sequence 1875, Ap
45	18.2	56.9	533	10	US-09-880-107-1875	

ALIGNMENTS

RESULT 1

US-09-070-927A-747/c

; Sequence 747, Application US/09070927A

; Patent No. US20020120116A1

; GENERAL INFORMATION:

; APPLICANT: Charles A. Kunsch

; Patrick J. Dillon

; Steven Barash

; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides

; NUMBER OF SEQUENCES: 982

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Human Genome Sciences, Inc.

; STREET: 9410 Key West Avenue

; CITY: Rockville

; STATE: Maryland

; COUNTRY: USA

; ZIP: 20850

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage

; COMPUTER: HP Vectra 486/33

; OPERATING SYSTEM: MSDOS version 6.2

; SOFTWARE: ASCII Text

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/070,927A

; FILING DATE: 04-May-2000

; CLASSIFICATION: <Unknown>

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 60/046,655

; FILING DATE: 1997-05-16

; APPLICATION NUMBER: 60/044,031

; FILING DATE: 1997-05-06

; APPLICATION NUMBER: 60/066,009

; FILING DATE: 1997-11-14

; ATTORNEY/AGENT INFORMATION:

; NAME: Kenley K. Hoover

; REGISTRATION NUMBER: 40,302

; REFERENCE/DOCKET NUMBER: PB369

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (301) 309-8504

; TELEFAX: (301) 309-8512

; INFORMATION FOR SEQ ID NO: 747:

; SEQUENCE CHARACTERISTICS:

```
;
; LENGTH: 1112 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 747:
US-09-070-927A-747

Query Match          66.9%; Score 21.4; DB 10; Length 1112;
Best Local Similarity 80.6%; Pred. No. 65;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATGTTGTTGTTATT 31
   || ||||| ||||| ||||| ||||| ||||| ||
Db 692 TAAACACACAAATATATATATATATGTTGTTT 662

RESULT 2
US-09-938-842A-4932
; Sequence 4932, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4932
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4932

Query Match          66.9%; Score 21.4; DB 9; Length 2000;
Best Local Similarity 80.6%; Pred. No. 64;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 AATACACACAAATATTAATGTTGTTGTTATTA 32
   || ||||| ||||| ||||| ||||| ||||| ||
Db 445 AAACACACAAATATTAATGTTGTTATGA 475

RESULT 3
US-09-938-842A-4407
; Sequence 4407, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
```

```
; SEQ ID NO 4407
; LENGTH: 1876
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4407

Query Match          65.0%; Score 20.8; DB 9; Length 1876;
Best Local Similarity 78.1%; Pred. No. 1e+02;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATGTTGTTGTTATTA 32
   ||||| ||||| ||||| ||||| ||||| |||||
Db 295 TAATACAAATCAATCTATTGTTGGATCAATTA 326

RESULT 4
US-09-764-847-1916/C
; Sequence 1916, Application US/09764847
; Patent No. US20020132767A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC009
; CURRENT APPLICATION NUMBER: US/09/764,847
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2003
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1916
; LENGTH: 18501
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (9022)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: SITE
; LOCATION: (9023)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-847-1916

Query Match          65.0%; Score 20.8; DB 10; Length 18501;
Best Local Similarity 78.1%; Pred. No. 97;
Matches 25; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATGTTGTTGTTATTA 32
   ||| ||||| ||||| ||||| ||||| ||||| ||
Db 17277 TAAATACACTAAACTAAATGTTGTTATTA 17246

RESULT 5
US-09-754-853A-2
; Sequence 2, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754,853A
; CURRENT FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 2
; LENGTH: 335913
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: CDS
```

```

: LOCATION: (45163)..(45314),(45450)..(45509),(46941)..(48763),(48975)..(49573)
: OTHER INFORMATION: Clone ID: 240017_region_G3
US-09-754-853A-2

```

Query Match 65.0%; Score 20.8; DB 9; Length 335913;
Best Local Similarity 78.1%;
Matches 25; Conservative 0.0; Mismatch 0.0;

Qy . 1 TAATACACACAAATATTAAATGTTGTGATTA 32
 ||||| | - ||||| | - ||||| |
 Db 311066 TAATACAAAATGAGATTAAATGGTTATGATTA 311097

RESULT 6

```

US-09-754-853A-3
: Sequence 3, Application US/09754853A
: Publication No. US20030005491A1
: GENERAL INFORMATION:
: APPLICANT: Hauge, Brian M.
: APPLICANT: Parnell, Laurence D.
: APPLICANT: Parsons, Jeremy D.
: APPLICANT: Wang, Ming Li
: TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
: TITLE OF INVENTION: Soybean Cyst Nematode Resistance
: FILE REFERENCE: 38-10(15810)B
: CURRENT APPLICATION NUMBER: US/09/754,853A
: CURRENT FILING DATE: 2001-01-05
: PRIOR APPLICATION NUMBER: US 60/174,880
: PRIOR FILING DATE: 2000-01-07
: NUMBER OF SEQ ID NOS: 1119
: SEQ ID NO 3
: LENGTH: 335913
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (46798)..(48975)..(49573)
: OTHER INFORMATION: Clone ID: 240017_region.G3
US-09-754-853A-3

```

Query Match 65.0%; Score 20.8; DB 9; Length 335913;
Best Local Similarity 78.1%;
Matches 25; Conservative 0.0; Mismatch 0.0;

Qy 1 TAATACAAACACAATATTAAATGGTGTGCTATTA 32
 ||||| | | ||||| | | |||||
 Db 311066 TAATACAAATGAGATTAATGGTTATGTATTA 311097

RESULT.7

```

US-09-764-860-770
; Sequence 770, Application US/09764860
; Patent No. US20020094953A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC008
; CURRENT APPLICATION NUMBER: US/09/764,860
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1198
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 770
; LENGTH: 504
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-860-770

Query Match 63.7% Score 20.4; DB 10; Length 504;
Best Local Similarity 80.0%; Pred. NO. 1.4e+02;
Matches 24; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```

QY 1 TAAATACACACAATATTAAATTGGTGTGTAT 30

Db 435 TAATGTAACACATTATTTCAGTGTCTTGTA 464

RESULT 8

```

US-09-960-352-11904/c
; Sequence 11904, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Machialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960, 352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11904
; LENGTH: 448
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB3057-016-Q1-K1-E4
US-09-960-352-11904

```

Query Match 63.1%; Score 20.2; DB 10; Length 448;
Best Local Similarity 88.0%; Pred. No. 1.6e+02;
Matches 22; Conservative 0; Mismatches 3; Indels 0

Qy	1	TAATACAACACAATATTAAATTGTGT	25
Db	289	TAATAAACACAATATTAAATGTAT	265

RESULT 9

```

US-09-917-800A-739/c
: Sequence 739, Application US/09917800A
: Patent No. US20020119462A1
: GENERAL INFORMATION:
: APPLICANT: Mendrick, Donna
: APPLICANT: Porter, Mark
: APPLICANT: Johnson, Kory
: APPLICANT: Castle, Arthur
: APPLICANT: Elashoff, Michael
: APPLICANT: Gene Logic, Inc.
: TITLE OF INVENTION: Molecular
: Toxicology Modeling
: FILE REFERENCE: 44921-5038-US
: CURRENT APPLICATION NUMBER: US/09/917,800A
: CURRENT FILING DATE: 2001-07-31
: PRIOR APPLICATION NUMBER: US 60/222,040
: PRIOR FILING DATE: 2000-07-31
: PRIOR APPLICATION NUMBER: US 60/222,880
: PRIOR FILING DATE: 2000-11-02
: PRIOR APPLICATION NUMBER: US 60/290,029
: PRIOR FILING DATE: 2001-05-11
: PRIOR APPLICATION NUMBER: US 60/290,645
: PRIOR FILING DATE: 2001-05-15
: PRIOR APPLICATION NUMBER: US 60/292,336
: PRIOR FILING DATE: 2001-05-22
: PRIOR APPLICATION NUMBER: US 60/295,798
: PRIOR FILING DATE: 2001-06-06
: PRIOR APPLICATION NUMBER: US 60/297,457
: PRIOR FILING DATE: 2001-06-13
: PRIOR APPLICATION NUMBER: US 60/298,884
: PRIOR FILING DATE: 2001-06-19
: PRIOR APPLICATION NUMBER: US 60/303,459
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 1740
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 739
: LENGTH: 562
: TYPE: DNA

```

; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 AI102871

US-09-917-800A-739

Query Match 61.9%; Score 19.8; DB 10; Length 562;
Best Local Similarity 77.4%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAATACAAACAATAATTGAATTTGCTTGTATT 31
 ||| ||| ||||| |
Db 451 TAGTCARAAAACAATAATTGAATTTGCCCTTAAT 421

RESULT 10
US-10-098-841-217
Sequence 217, Application US/10098841
Publication NO. US20020197679A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi Vinod
APPLICANT: Xu, Chongjun
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunqing
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhao, Qing A.
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Wehrman, Tom
APPLICANT: Zhang, Jie
APPLICANT: Qian, Xiaohong B.
APPLICANT: Dumanac, Radje T.

TITLE OF INVENTION: No. US20020197679A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides

FILE REFERENCE: 784CIP2

CURRENT APPLICATION NUMBER: US/10/098,841

PRIOR FILING DATE: 2002-03-13

PRIOR APPLICATION NUMBER: 09/598,042

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 331

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 217

LENGTH: 1936

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (81)..(752)

NAME/KEY: misc.feature

LOCATION: (1)...(1936)

OTHER INFORMATION: n = a,t,c or g

US-10-098-841-217

Query Match 61.9%; Score 19.8; DB 9; Length 1936;
Best Local Similarity 77.4%; Pred. No. 2.2e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 TAATACAAACAATAATTGAATTTGCTTGTATT 31
 ||| ||| ||||| |
Db 1448 TAATCAATCAATAATTGAATTTGCCTTAAT 1478

RESULT 11
US-09-984-245-78
Sequence 78, Application US/09984245
Patent No. US20020165374A1

```

RESULT 12
US-09-754-853A-1
; Sequence 1, Application US/09754853A
; Publication No. US20030005491A1
; GENERAL INFORMATION:
; APPLICANT: Hauge, Brian M.
; APPLICANT: Parnell, Laurence D.
; APPLICANT: Parsons, Jeremy D.
; APPLICANT: Wang, Ming Li
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-10(15810)B
; CURRENT APPLICATION NUMBER: US/09/754.853A
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 60/174,880
; PRIOR FILING DATE: 2000-01-07
; NUMBER OF SEQ ID NOS: 1119
; SEQ ID NO 1
; LENGTH: 127197
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 515002_region_G2
US-09-754-853A-1

```

```

Query Match      61.9%; Score 19.8; DB 9; Length 127197;
Best Local Similarity 77.4%; Pred. No. 1.9e+02;
Matches 24; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

Qy 1 TAATACACACAAATTAATTGTTGTGTTAT 31
    ||||| ||||| ||| ||| ||||| |||
Db 93703 TAATACACAGAATCGTGTGTTTGTGTTGTT 93733

```

```

RESULT 13
US-10-254-869-3/c
; Sequence 3, Application US/10254869
; Publication No. US20030027307A1
; GENERAL INFORMATION:
; APPLICANT: YE, Jane et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01160DIV
; CURRENT APPLICATION NUMBER: US/10/254,869
; CURRENT FILING DATE: 2002-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 148567
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(148567)
; OTHER INFORMATION: n = A,T,C or G
US-10-254-869-3

```

Query Match	61.9%	Score 19.8;	DB 9;	Length 148557;
Best Local Similarity	77.4%;	Prd. No. 1.9e+02;		
Matches 24;	Conservative 0;	Mismatches 7;	Indels 0;	Gaps 0;
QY	2	ANTACACACAAATTAATTGTTGTATTATTA	32	
DB	1241180	ACACAGACAGACATGAATTGTTGTATTATTA	1241150	

RESULT 14
US-09-801-876B-3/c
; Sequence 3, Application US/09801876B
; Patent No. US20020127683A1

```

: GENERAL INFORMATION:
: APPLICANT: YE, Jane et al
: TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
: TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
: TITLE OF INVENTION: THEREOF
: FILE REFERENCE: CLO01160
: CURRENT APPLICATION NUMBER: US/09/801.876B
: CURRENT FILING DATE: 2001-03-09
: NUMBER OF SEQ ID NOS: 8
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 3
: LENGTH: 148567
: TYPE: DNA
: ORGANISM: Human
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)..(148567)
: OTHER INFORMATION: n = A, T, C or G
: US-09-801-876B-3

```

```

Query Match      61.9%   Score 19.8;   DB 10;
Best Local Similarity 77.4%   Pred. No. 1.9e+02;
Matches 24;   Conservative 0;   Mismatches 7;   Indels 0;   Gaps 0;

QY      2  ATATACACACAATATTAATGTGTGTGTATTA  32
      || | | | | | | | | | | | | | | | |
Db 124180 AACAGACAGACAGACGAATGTGTGTATTA  124150

```

```

RESULT 15
US-09-770-791-596
; Sequence 596, Application US/09770791
; Patent No. US20020062014A1
; GENERAL INFORMATION:
; APPLICANT: Goriach, Jörn
; APPLICANT: An, Yong-Qiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Kricker, Maja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hurban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis thaliana
; FILE REFERENCE: 2029 (PARA-018PRV)
; CURRENT APPLICATION NUMBER: US/09/770,791
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 60/178,480
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 596
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-791-596

```

Query Match 61.3%; Score 19.6; DB 10; Length 357;
Best Local Similarity 84.8%; Pred. No. 2.6e+02;
Matches 22; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
1 TAATACACACAAATATTAAATTGTGT 26

us-09-895-435-3.rnpb

Thu Feb 13 12:34:56 2003

Db 18 TAATTCACACAAATATTCATCGTATT 43

Search completed: February 13, 2003, 04:56:51
Job time : 241 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run On: February 12, 2003, 23:34:41 : Search time 2255 Seconds
(without alignments)
229.825 Million cell updates/sec

Title: US-09-895-435-3
Perfect score: 32
Sequence: 1 taataacacacataattgtgtgttattta 32

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST :
1: em_estba :
2: em_esthum :
3: em_estin :
4: em_estnu :
5: em_estov :
6: em_estpl :
7: em_estro :
8: em_htc :
9: gb_est1 :
10: gb_est2 :
11: gb_htc :
12: gb_est3 :
13: gb_est4 :
14: gb_est5 :
15: em_estfun :
16: em_estom :
17: gb_gss :
18: em_gss_hum :
19: em_gss_inv :
20: em_gss_pln :
21: em_gss_vrt :
22: em_gss_fun :
23: em_gss_mam :
24: em_gss_mus :
25: em_gss_other :
26: em_gss_pro :
27: em_gss_rod :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
c 1	24	75.0	953	17	CNS0431LN
c 2	23.6	73.8	510	17	AZ523535 220PBF09
c 3	23.2	72.5	700	17	AJ227574 Ciona int
c 4	23	71.9	459	9	AA677944 z124h07.s
5	22.8	71.2	1101	17	CNS0002SR
6	22.6	70.6	526	9	AA101244 zn43b06.s

7	22.4	70.0	523	17	AQ970499
c 8	22.4	70.0	562	17	AZ848763
c 9	22.4	70.0	584	14	BQ692893
c 10	22.4	70.0	633	17	AQ450340
11	22	68.8	239	13	BJ365887
12	22	68.8	505	10	AV385246
13	22	68.8	571	17	AQ319355
c 14	21.8	68.1	1042	17	CNS05J05
15	21.6	67.5	536	14	BQ693146
16	21.4	66.9	554	10	AW334296
c 17	21.4	66.9	610	17	BH762639
c 18	21.4	66.9	716	17	BH536105
c 19	21.4	66.9	808	17	BH452731
c 20	21.4	66.9	810	17	BH419970
c 21	21.4	66.9	831	17	BH589798
c 22	21.4	66.9	843	17	BH248522
23	21.4	66.9	850	17	BH161606
24	21.4	66.9	913	17	BH157595
c 25	21.4	66.9	920	17	BH134609
c 26	21.4	66.9	958	17	B08958
c 27	21.4	66.9	1154	17	CNS06PT8
28	21	65.6	345	9	AA860018
c 29	21	65.6	484	17	BH741995
30	21	65.6	510	9	A1726715
31	21	65.6	583	9	A1729966
c 32	21	65.6	603	17	PT022K01R
c 33	21	65.6	603	17	AQ662080
c 34	21	65.6	626	9	A1730796
35	21	65.6	630	17	AZ883204
c 36	21	65.6	684	17	BH698214
37	21	65.6	831	12	BC441984
38	21	65.6	1007	17	CNS02W13
39	21	65.6	1070	17	CNS03PL2
40	21	65.6	1130	12	BG033546
41	20.8	65.0	281	13	BI889362
42	20.8	65.0	303	17	AQ005284
c 43	20.8	65.0	318	17	B54164
c 44	20.8	65.0	342	10	AV559105
c 45	20.8	65.0	402	10	AW464877

ALIGNMENTS

RESULT 1	CNS0431LN/c	953 bp	DNA	linear	GSS 18-MAY-2000
LOCUS	Tetraodon nigroviridis genome survey sequence PUC-ori end of clone 079N02 of library G from Tetraodon nigroviridis, genomic survey sequence.				
DEFINITION	AL272948				
ACCESSION	AL272948.1	GI:7995193			
VERSION	GSS: genome survey sequence.				
KEYWORDS	Tetraodon nigroviridis.				
SOURCE	Tetraodon nigroviridis				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.				
REFERENCE	1 (bases 1 to 953)				
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.				
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 953)				
AUTHORS	Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.				
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis				
JOURNAL	Unpublished				

```

REFERENCE
AUTHORS      3 (bases 1 to 953)
TITLE        Genoscope.
JOURNAL      Direct Submission
COMMENT      Submitted (12-APR-2000)
              This sequence is a single read and was generated as part of a large
              scale clone-end sequencing project of the Tetraodon nigroviridis
              genome. For more information, please take a look at
              http://www.genoscope.cns.fr/tetraodon.
FEATURES
source
  1..953
    /organism="Tetraodon nigroviridis"
    /db_xref="taxon:99883"
    /clone="079N02"
    /clone_lib="G"
    /note="Genoscope sequence ID : COBG079DG01SP1-end :
    PUC-Ori"
BASE COUNT   316 a 218 c 176 g 238 t 5 others
ORIGIN
Query Match 75.0%; Score 24; DB 17; Length 953;
Best Local Similarity 84.4%; Pred. No. 98;
Matches 27; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATTGTTGTTGTTATTA 32
    ||||| ||| ||||| ||||| ||||| |||||
Db 233 TAATAAACTCAATATTTATTTGGGTTTATTA 202

RESULT 2
AZ523535/c
LOCUS       AZ523535
DEFINITION 220PbF09 Pb MBN #21 Plasmodium berghei genomic 3', DNA sequence.
ACCESSION  AZ523535
VERSION     AZ523535.1 GI:13962603
KEYWORDS   GSS.
SOURCE     Plasmodium berghei.
ORGANISM   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE  1 (bases 1 to 510)
AUTHORS    Carlton,J.M.-R. and Dame,J.B.
TITLE      The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL    Parasitol Today (Regul. Ed.) 16 (10), 409 (2000)
COMMENT    Contact: Dame JB
            Dept. of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: damejb@mail.vetmed.ufl.edu
            Seq primer: M13(-20) forward
            Class: shotgun.
FEATURES
source
  1..510
    /organism="Plasmodium berghei"
    /strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
    /db_xref="taxon:5821"
    /clone_lib="Pb MBN #21"
    /dev_stage="asexual blood forms"
    /lab_host="Mus musculus"
    /note="vector: pBluescript SK(+); vector DNA, phagemid
    excised from lambda ZAP; Site_1: EcoRV; Site_2: EcoRV;
    Genomic DNA was prepared from asynchronous blood stage
    forms of the cloned ANKA isolate of P. berghei grown in
    laboratory Swiss white mice. The DNA was purified from
    contaminating host DNA by Hoechst Dye 33258-CsCl
    ultracentrifugation and precipitated. Purified DNA was
    digested withmung bean nuclease in the presence of 36-38%
    formamide at 50 C, as described (Vernick, K.D., Imberski,
    R.B., and McCutchan, T.F. 1988. Nucleic Acids Research
    16:6883-6896). The ends of the digestion fragments were
    polished using T4 DNA polymerase, and the fragments size
    selected in the range 500-2000 bp. These were ligated into
    the EcoRV-cleaved and dephosphorylated pBluescript SK(+)
```

```

vector. Recombinant plasmids were used to transform E.
coli XL10-Gold host cells.
BASE COUNT   177 a 65 c 42 g 226 t
ORIGIN
Query Match 73.8%; Score 23.6; DB 17; Length 510;
Best Local Similarity 86.7%; Pred. No. 1.4e+02;
Matches 26; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATTGTTGTTGTTAT 30
    ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TAATAAACAAATAATAATAATTGTTGTTTAT 473

RESULT 3
CI9D10/c
LOCUS       CI9D10
DEFINITION Ciona intestinalis genomic fragment, clone 9D10, genomic survey
sequence.
ACCESSION  AJ227574
VERSION     AJ227574.1 GI:2951400
KEYWORDS   GSS; genome survey sequence.
SOURCE     Ciona intestinalis.
ORGANISM   Ciona intestinalis.
            Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
            Phlebobranchia; Cionidae; Ciona.
REFERENCE  1 (bases 1 to 700)
AUTHORS    Simmen,M.W., Leitgeb,S., Clark,V.H., Jones,S.J. and Bird,A.
TITLE      Gene number in an invertebrate chordate, Ciona intestinalis
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 95 (8), 4437-4440 (1998)
MEDLINE    98208558
PUBMED     9539755
REFERENCE  2 (bases 1 to 700)
AUTHORS    Simmen,M.W., Leitgeb,S., Charlton,J., Jones,S.J., Harris,B.R.,
            Clark,V.H. and Bird,A.
TITLE      Nonmethylated transposable elements and methylated genes in a
            chordate genome
JOURNAL    Science 283 (5405), 1164-1167 (1999)
MEDLINE    99148102
PUBMED     10024242
REFERENCE  3 (bases 1 to 700)
AUTHORS    Simmen,M.W. and Bird,A.
TITLE      Sequence analysis of transposable elements in the sea squirt, Ciona
            intestinalis
JOURNAL    Mol. Biol. Evol. 17 (11), 1685-1694 (2000)
MEDLINE    20523971
PUBMED     11070056
REFERENCE  4 (bases 1 to 700)
AUTHORS    Clark,V.H., Leitgeb,S., Charlton,J., Tweedie,S., Simmen,M.W. and
            Bird,A.P.
TITLE      Direct Submission
JOURNAL    Submitted (28-FEB-1998) ICMB, University of Edinburgh, King's
            Buildings, Mayfield Rd, Edinburgh EH9 3JR, UK. E-mail contact:
            VCLARK@sv0.bio.ed.ac.uk
            Vector: pBluescript KS.
FEATURES
source
  1..700
    /organism="Ciona intestinalis"
    /db_xref="taxon:7719"
    /clone="9D10"
    /dev_stage="adult"
BASE COUNT   251 a 105 c 115 g 229 t
ORIGIN
Query Match 72.5%; Score 23.2; DB 17; Length 700;
Best Local Similarity 89.3%; Pred. No. 1.9e+02;
Matches 25; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 5 ACAACACAAATTAATTGTTGTTGTTATTA 32
    ||||| ||||| ||||| ||||| ||||| |||||
Db 673 ACAACACAGATATAAATTGTTGTTATTA 646
```

```

RESULT 4
AA677944/c
LOCUS
DEFINITION
  z124h07.s1 Soares_fetal_liver_spleen_INFLS_S1 Homo sapiens cDNA
  clone IMAGE:431773 3', mRNA sequence.
ACCESSION
AA677944
VERSION
AA677944.1 GI:2658466
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 459)
Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 451.
FEATURES
Location/Qualifiers
1..459
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:431773"
/clone_lib="Soares_fetal_liver_spleen_INFLS_S1"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
with a modified polylinker; Site.1: Pac I; Site.2: Eco RI;
This is a subtracted version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
ACTGCGAGCAATTAAATTAAGATCTTTTTTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT
145 a 82 c 102 g 129 t 1 others
ORIGIN
Query Match 71.9%; Score 23; DB 9; Length 459;
Best Local Similarity 83.9%; Pred. No. 2.3e+02;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 2 AATACACACAAATATTAATGCTGTTTATTA 32
|||||
Db 410 AATATAACACAAATATTAATGCTGTTTATGA 380

RESULT 5
CNS002SR
LOCUS
DEFINITION
  Drosophila melanogaster genome survey sequence TET3 end of BAC #
  BACR07G08 of RPCT-98 library from Drosophila melanogaster (fruit
  fly), genomic survey sequence.
ACCESSION
AL063320
VERSION
AL063320.1 GI:4941177
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster.
ORGANISM
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

RESULT 6
AA101244
LOCUS
DEFINITION
  zn43b06.s1 Stratagene HeLa cell s3 937216 Homo sapiens cDNA clone
  IMAGE:550163 3', mRNA sequence.
ACCESSION
AA101244
VERSION
AA101244.1 GI:1647884
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 526)
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
Schellenberg,K., Soares,M.B., Tan,F., Thierry-Mieg,J., Trevaskis,E.,
Underwood,K., Wohlmann,P., Waterston,R., Wilson,R. and Marra,M.
Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40M13 fwd. from Amersham

```

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1101)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Hammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCT-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
Source

BASE COUNT 248 a 214 c 174 g 374 t 91 others
ORIGIN

Query Match 71.2%; Score 22.8; DB 17; Length 1101;
Best Local Similarity 65.6%; Pred. No. 2.6e+02;
Matches 21; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

QY 1 TAATACACACAAATATTAATGCTGTTTATTA 32
|||||
Db 940 TAATAACACAAATATTAATGCTGTTTATTTATTT 971

RESULT 6

AA101244

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

COMMENT

High quality sequence stop: 233.

FEATURES

Location/Qualifiers
1. .526
/organism="Homo sapiens"
/db_xref="GDB:3928681"
/db_xref="taxon:9606"
/clone="IMAGE:550163"
/clone_lib="Stratagene HeLa cell s3 937216"
/sex="female"
/dev_stage="HeLa S3 cell line"
/lab_host="SOLR (kanamycin resistant)"
/note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer: Oligo dT. HeLa S3 epitheloid carcinoma cells grown to semi-confluency without induction. Average insert size: 1.5 kb; Uni-ZAP XR Vector. -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTT 3"

BASE COUNT 202 a 94 c 57 g 165 t 8 others
ORIGIN

Query Match 70.6%; Score 22.6; DB 9; Length 526;
Best Local Similarity 83.3%; Pred. No. 3.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TAATACACACATATTAATTTGGTTGTAT 30
||||||| ||||||| ||||||| |||
Db 390 TATACACGCAATATTAATTTGGTTTAT 419

RESULT 7
AQ970499
LOCUS
DEFINITION
RPCI-23-325C1.TJ RPCI-23 Mus musculus genomic clone RPCI-23-325C1, DNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P., and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-325C1.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@edlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
Plate: 325 row: C column: 1
Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
1. .523
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-325C1"
/clone_lib="RPCI-23"
/sex="Female"
/lab_host="DH10B"

FEATURES

Location/Qualifiers
1. .562
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0149H21"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, P-"
/note="Vector: PWB42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site.1: EcoRI; Site.2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI. MethyIase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 173 a 109 c 120 g 120 t 1 others
ORIGIN

Query Match 70.0%; Score 22.4; DB 17; Length 523;
Best Local Similarity 81.2%; Pred. No. 3.7e+02;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 TAATACACACATATTAATTTGGTTGTATTA 32
||||||| ||||||| ||||||| |||

Db 70 TATACAAATATTAATTTGGTTGTATTA 101

RESULT 8

AZ848763/C
LOCUS
DEFINITION
2M0149H21R Mouse 10kb plasmid UUCG1M library Mus musculus genomic clone UUCG2M0149H21 R, DNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0149 row: H column: 21
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 562.
Location/Qualifiers
1. .562
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0149H21"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, P-"
/note="Vector: PWB42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative

FEATURES

Location/Qualifiers
1. .562
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUCG2M0149H21"
/clone_lib="Mouse 10kb plasmid UUCG1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, P-"
/note="Vector: PWB42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative

```

Query Match      70.0%; Score 22.4; DB 14; Length 584;
Best Local Similarity 81.2%; Pred. No. 3.7e+02;
Matches 28; Conservative 0; Mismatches 6; Indels 0; Caps 0;

QY 1 TAATACACACAATATTAATTTGTTGTATTA 32
    |||||
Db 148 TATTATTAATTTATATTAATTTGTTGTATTA 117

```

RESULT_10					
AQ450340/c					
LOCUS	AQ450340	633 bp	DNA	linear	GSS 08-APR-1999
DEFINITION	500010F04.x1 CpIOWAM13mpl8gNAL Cryptosporidium parvum genomic, DNA sequence.				

LOCUS	AA043040	633 bp	DNA	linear	GSS 08-APR-1999
DEFINITION	500010F04.x1 CpIOWAM1.3mpl8gDNA1 Cryptosporidium parvum genomic, DNA sequence.				
ACCESSION	AA043040				
VERSION	GSS				
KEYWORDS	AA050340.1 GI:4579477				
SOURCE	Cryptosporidium parvum.				
ORGANISM	Cryptosporidium parvum.				

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Alveolata; Apicomplexa; Coccidia; Elmeriida;
Cryptosporidiidae; Cryptosporidium.
1 (bases 1 to 633)
Hyman, R.W., Fung, E., Qin, F., Rowley, D. and Davis, R.W.
Cryptosporidium parvum genome sequencing demonstration project
Unpublished (1999)
Contact: Hyman, R.W.

TITLE	Cryptosporidium parvum genome sequencing demonstration project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Hyman, R. W. Stanford DNA Sequencing and Technology Center Stanford University School of Medicine, Palo Alto 855 California Avenue, Palo Alto, CA 94304, USA Tel: 650 812 1972 Fax: 650 812 1975

Email: hyman@sequence.stanford.edu
For Annotation Data see <http://medsfgh.ucsf.edu/id/CpTags/home.html>
Seq primer: M13(-21) forward
Class: shotgun.

```

FEATURES
Source
Location/Qualifiers
1. 533
/organism="Cryptosporidium parvum"
/strain="IOWA"
/db_xref="taxon:5807"
/clone_lib="CpIOWAM13mp18gDNa1"
/lab_host="E. coli DH125"

```

/note="Vector: M13mp18; Site_1: Hind III; C. parvum (TOWA isolate) genomic DNA was hydrodynamically sheared to produce fragments having a tight size distribution between 1.5 and 3 kb. Adaptors (pGT10ACTCA/CAAAACCCTAGTTP) were ligated to the randomly sheared gDNA fragments and pAGCTGTTT linkers were ligated to the Hind III-cleaved M13mp18 vector. The adaptor-containing inserts were annealed and ligated to the vector and transformed into E. coli strain DH125. Recombinant phagemid clones from the first plating of the library were randomly selected for sequence analysis using the M13(-21) forward primer."

```

SEQUENCE ANALYSIS USING THE M13(-21) FORWARD PRIMER."
236 a      77 c    101 g     219 t
ORIGIN
Query Match          70.0%; Score 22.4; DB 17; Length 633;
Best Local Similarity 81.2%; Pred. No. 3.7e+02;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
1 TAATACACACAATAATTAAATGTTGTGTATTGA 32
T | G | I | I | I | I | I | I | I | I | I | I | I | I | I |
bb 337 TTATGCAAAACAATTTTAATTTGGCGCTGTATTA 306

RESULT 11
BU365887 LOCUS
BU365887 DEFINITION
Dictyostelium discoideum cDNA library, Cf Dictyostelium
discoideum cDNA clone ddc36n24 5', mRNA sequence.
BU365887 ACCESSION
BU365887 VERSION
BU365887.1 GI:19275189
EST 08-MAR-2002

```

LOCUS	Accession	Length	Library	EST
LOCUS	U00001	1329 bp	linear	U00001
DEFINITION	U00001	1329 bp	linear	U00001
Accession	U00001	1329 bp	linear	U00001
Version	U00001	1329 bp	linear	U00001

Thu Feb 13 12:34:57 2003

```

KEYWORDS
SOURCE Dictyostelium discoideum.
ORGANISM Dictyostelium discoideum
REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
AUTHORS 1 (bases 1 to 239)
TITLE Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
JOURNAL Full length cDNA of Dictyostelium discoideum at the culmination
COMMENT stage
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
Location/Qualifiers
1..239
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/clone="ddc36n24"
/clone_lib="Dictyostelium discoideum cDNA library, CF"
/sex="mat A"
/dev_stage="Culmination stage"
BASE COUNT 113 a 26 c 1 g 99 t
ORIGIN
Query Match 68.8%; Score 22; DB 13; Length 239;
Best Local Similarity 83.3%; Pred. No. 5.3e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 2 AATAACAACAATTAATTTGTGTGTTAT 31
|||||
Db 92 AATAACAACAATTAATTTATTTTCTT 121

RESULT 12
AV385246 505 bp mRNA linear EST 27-OCT-1999
LOCUS AV385246 Halocynthia roretzi Fertilized egg Halocynthia roretzi
DEFINITION cDNA clone 005A08_3 3', mRNA sequence.
ACCESSION AV385246
VERSION AV385246.1 GI:6130303
KEYWORDS EST.
SOURCE Halocynthia roretzi.
ORGANISM Halocynthia roretzi
REFERENCE Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea;
AUTHORS Stolidobranchia; Pyuridae; Halocynthia.
TITLE 1 (bases 1 to 505)
JOURNAL Makabe,K.W.
COMMENT Halocynthia roretzi EST
Contact: Kazuhiro W. Makabe
Department of Zoology, Graduate School of Science
Kyoto University
Sakyo-ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4095
Fax: 81-75-705-1113
Email: kwmakabe@scs.kyoto-u.ac.jp.
FEATURES
source
Location/Qualifiers
1..505
/organism="Halocynthia roretzi"
/db_xref="taxon:7729"
/clone="005A08_3"
/clone_lib="Halocynthia roretzi Fertilized egg"
/dev_stage="Fertilized egg"
/note="Organ: embryo"
BASE COUNT 175 a 79 c 65 g 186 t
ORIGIN
Query Match 68.8%; Score 22; DB 10; Length 505;
Best Local Similarity 83.3%; Pred. NO. 5.2e+02;

Qy 3 ATACAACAATTAATTTGTGTGTTATTA 32
|||||
Db 313 ATTCAACAGATATTTATTTATATTA 342

RESULT 14
CNS05J05 1042 bp DNA linear GSS 26-MAY-2000
LOCUS CNS05J05/c Tetradon nigroviridis genome survey sequence SP6 end of clone
DEFINITION 017F15 of library B from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL340430
VERSION AL340430.1 GI:8234188
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.

```

```

Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 TAATACAACAATTAATTTGTGTGTTAT 30
|||||
Db 299 TCAACATAACAATTAATTTGTGTATAT 328

RESULT 13
AQ319355 571 bp DNA linear GSS 06-MAY-1999
LOCUS RPC111-97J9.TJ RPC1-11 Homo sapiens genomic clone RPC1-11-97J9, DNA
DEFINITION sequence.
ACCESSION AQ319355
VERSION AQ319355.1 GI:4052320
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 571)
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,
Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.
Use of human BAC End Sequences for Sequence-Ready Map Building
Unpublished (1998)
Other_GSSs: RPC111-97J9.TJ
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPC1-11. For BAC
library availability, please contact pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/hungen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
source
Location/Qualifiers
1..571
/organism="Homo sapiens"
/db_xref="GDB:7537088"
/db_xref="taxon:9606"
/clone="RPC1-11-97J9"
/clone_lib="RPC1-11"
/sex="Male"
/cell_type="Lymphocytes"
/note="Vector: pBACE3.6; Site.1: ECORI; Site.2: ECORI;
RPC111 Human Male BAC Library"
BASE COUNT 197 a 88 c 100 g 186 t
ORIGIN
Query Match 68.8%; Score 22; DB 17; Length 571;
Best Local Similarity 83.3%; Pred. No. 5.2e+02;
Matches 25; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 3 ATACAACAATTAATTTGTGTGTTATTA 32
|||||
Db 313 ATTCAACAGATATTTATTTATATTA 342

RESULT 14
CNS05J05/c 1042 bp DNA linear GSS 26-MAY-2000
LOCUS CNS05J05 Tetradon nigroviridis genome survey sequence SP6 end of clone
DEFINITION 017F15 of library B from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL340430
VERSION AL340430.1 GI:8234188
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.

```

ORGANISM	Tetraodon nigroviridis
REFERENCE	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
AUTHORS	1 (bases 1 to 1042) Roest-Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE	Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1042)
AUTHORS	Roest-Crolius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 1042)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000)
COMMENT	This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .

```

source      1. .1042
/organism="Tetraodon nigriviridis"
/db_xref="taxon:99883"
/clone="017F15"
/clone_lib="B"
/note="Genoscope sequence ID : C0AB017CC08B1-end : SP6"
297 a      213 c      219 g      282 t      31 others
BASE COUNT
ORIGIN

Query Match      58.1%; Score 21.8; DB 17; Length 1042;
Best Local Similarity 92.0%; Pred. No. 5.9e+02;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      7 AACACAATATTAATTCGTTGTATT 31
||||| |||| |||||||||
Db      71 AACACAATATTTCGTTGTATT 47

```

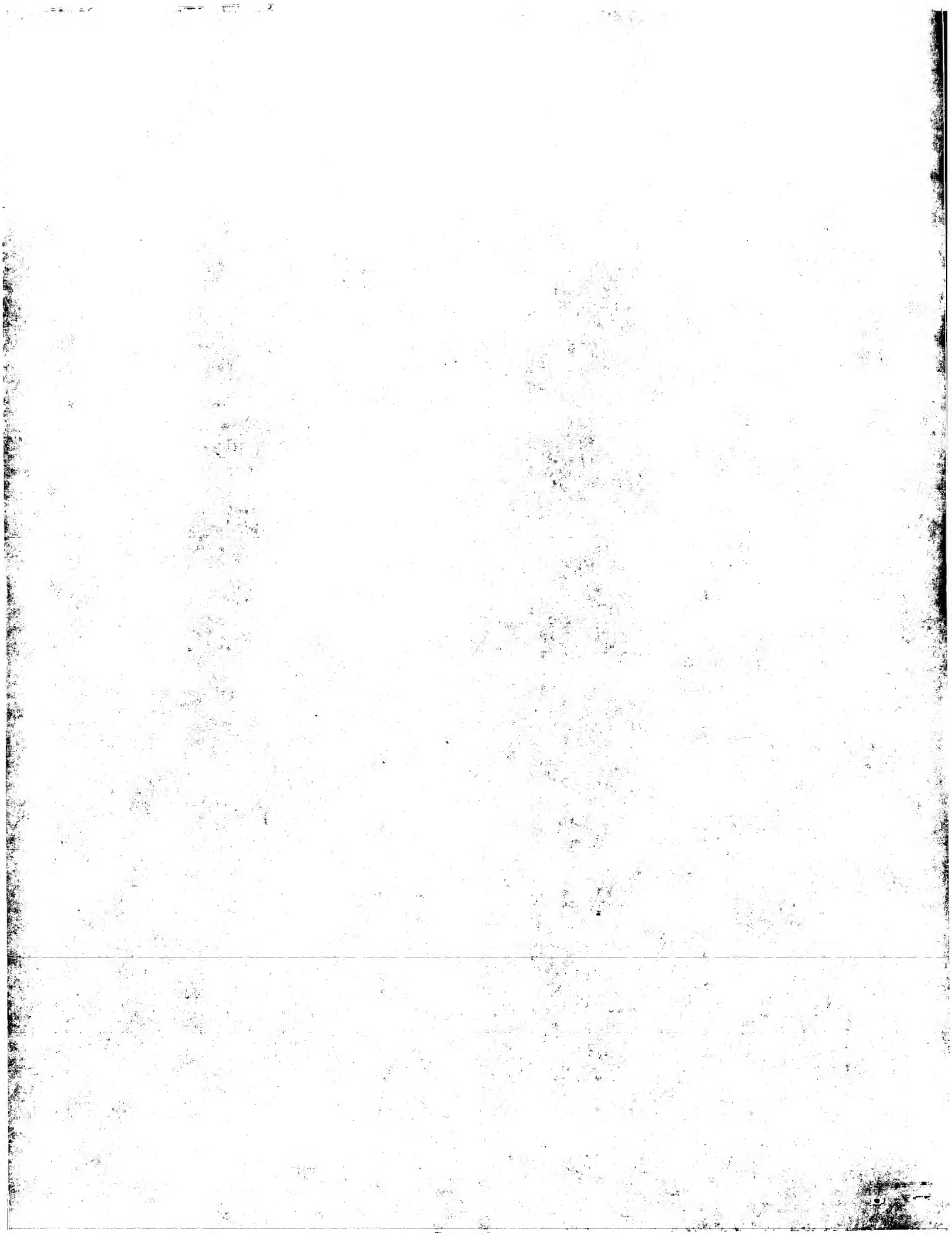
```

Tel: 314 286 1800
Fax: 314 286 1810
Email: estewartson.wustl.edu
The library was constructed by Claire Murphy and Dr. James McCarter
at Washington University, St. Louis. Newborn larvae were collected
from infected rats and provided by Dr. Judith Appleton of Cornell
University, Ithaca, NY (jaa2@cornell.edu).
Seq primer: -40RP from Gibco
High quality sequence stop: 414.
Location/Qualifiers
1. .536
/organism="Trichinella spiralis"
/db_xref="taxon:6334"
/clone_lib="Trichinella spiralis immature L1 pamp1 v1"
/dev_stage="immature L1"
/lab_host="DH10B"
/note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
The library was constructed by Claire Murphy and Dr. James
McCarter at Washington University, St. Louis. The cDNA was
made by using Dynabead oligo-dT priming (Dynal). PCR based
library using a modified protocol from the SMART PCR cDNA
Synthesis Kit from Clontech. Directionally cloned into the
UDG sites of pAMP1. Newborn larvae were collected from
infected rats and provided by Dr. Judith Appleton of
Cornell University, Ithaca, NY (jaa2@cornell.edu)."
```

BASE COUNT	171 a	85 c	110 g	170 t
ORIGIN				
Query Match	67.5%; Score 21.6; DB 14; Length 536;			
Best Local Similarity	85.7%; Pred. No. 7.2e+02;			
Matches	24;	Conservative	0;	Mismatches 4; Indels 0; Gaps 0;

Qy	4	TACAACACAATATTAAATTGTTGTATT	31
Db	411	TATAGGCATATTAAATGTTGTATT	438

Search completed: February 13, 2003, 03:17:52
Job time : 2259 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2003, 01:45:51 ; Search time 298 Seconds
(without alignments)
1881.703 Million cell updates/sec

Title: US-09-895-435-4
Perfect score: 249
Sequence: 1 ggggtaccgcagcatttgcg.....gacgcggaggaaatcacatg 249

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2185239 seqs, 1125999159 residues

Word size : 0
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	38	15.3	5216	AA07548	PKS90 TIR coupling
2	38	15.3	5230	AA07549	PTREX1 mutant beta
3	38	15.3	5230	AA015041	Nucleotide sequenc
4	38	15.3	5231	AA07550	PTREX1A mutant bet
5	38	15.3	5497	AA029646	Plasmid pTmTpf1.
6	38	15.3	5770	AA015043	Nucleotide sequenc
7	38	15.3	5870	AA015044	Nucleotide sequenc
8	38	15.3	5906	AA015042	Nucleotide sequenc
9	19	7.6	1566	AA013472	Enterococcus faeca

C 10	19	7.6	6076	22	AA046664	Tumour suppressor
C 11	19	7.6	6136	24	ABK40042	Human chemically p
C 12	19	7.6	7008	22	AA046531	Tumour suppressor
C 13	19	7.6	140036	24	AA098600	Human genomic DNA
C 14	18	7.2	251	21	AA022654	Human secreted pro
C 15	18	7.2	286	24	ABN16703	Human ORFX polynuc
C 16	18	7.2	342	22	AAH69737	Human cervical can
C 17	18	7.2	345	22	AAH73166	Human cervical can
C 18	18	7.2	457	22	ABA44883	Human foetal cell
C 19	18	7.2	457	22	ABA55346	Human foetal liver
C 20	18	7.2	457	22	ABA25077	Human breast cancer
C 21	18	7.2	457	22	AAK03588	Human brain expres
C 22	18	7.2	457	22	AAK29048	Human bone marrow
C 23	18	7.2	457	22	AAI13636	Probe #3569 for ge
C 24	18	7.2	457	22	AAI34995	Probe #3681 used t
C 25	18	7.2	457	22	AAI03520	Probe #3511 used t
C 26	18	7.2	457	24	ABS03593	Human genome-deriv
C 27	18	7.2	474	16	AAT19891	Human gene signatu
C 28	18	7.2	514	22	AAH70477	Human cervical can
C 29	18	7.2	557	22	ABL11160	Human cDNA clone (
C 30	18	7.2	602	23	ABL13041	Drosophila melanog
C 31	18	7.2	913	22	ABL23181	Human breast cance
C 32	18	7.2	1213	24	ABL55491	Alpha-catenin 10.3
C 33	18	7.2	2246	24	ABK54129	cDNA encoding huma
C 34	18	7.2	2561	23	ABL02994	Drosophila melanog
C 35	18	7.2	2582	23	ABL24840	Drosophila melanog
C 36	18	7.2	2642	22	AAH17727	Human cDNA sequenc
C 37	18	7.2	2673	21	AAH76657	Human ORFX ORF212
C 38	18	7.2	2695	23	ABL27234	Drosophila melanog
C 39	18	7.2	2717	23	ABL13040	Drosophila melanog
C 40	18	7.2	2733	22	AAH33639	Human colon cancer
C 41	18	7.2	2947	22	AAH44832	Rat cDNA encoding
C 42	18	7.2	3181	22	ABA08899	Human secreted pro
C 43	18	7.2	4704	22	AAI36927	Human musculoskele
C 44	18	7.2	4704	22	AAK67840	Human immune/haema
C 45	18	7.2	4704	22	AAK73415	Human immune/haema

ALIGNMENTS

RESULT 1
AA07548/c
ID AA07548 standard; DNA: 5216 BP.

XX AA07548;

XX 08-JUN-1999 (first entry)

XX PKS90 TIR coupling vector.

XX Beta-toxin; cpb; vaccine; infection; protection;

XX retained immunogenicity; haemorrhagic enteritis; necrotic enteritis;

XX enterotoxemia; lamb dysentery; translation initiation region; ds.

XX Synthetic.

XX EP892054-AL.

XX 20-JAN-1999.

XX 17-JUN-1998; 98EP-0202032.

XX 20-JUN-1997; 97EP-0201888.

XX (ALKU) AKZO NOBEL NV.

XX Frandsen PL, Sengers RPAM, Waterfield NR, Wells JM;

XX WPI; 1999-083571/08.

XX New detoxified derivative of Clostridium perfringens beta-toxin -

XX with retained immunogenicity, useful as a vaccine to protect against

PT Clostridium perfringens infection

XX Example; Fig 2b; 30pp; English.

XX The sequence is that of the vector pKS90 which was used in the

CC construction of a detoxified immunogenic derivative of Clostridium

CC perfringens beta-toxin. Such a derivative is useful as a vaccine

CC to induce an immune response and protect against Clostridium

CC perfringens infection in man, pig, lamb, sheep, goat, calf and

CC bird, which causes haemorrhagic enteritis, necrotic enteritis,

CC enterotoxemia and lamb dysentery. The new beta-toxin has been

CC detoxified without impairing the immunogenicity, unlike prior

CC art vaccines which used chemical and formalin detoxifying methods.

CC Production of the beta-toxin in gram positive bacteria prevents

CC prior art difficulties of isolation from dangerous Clostridium

CC perfringens, and purification from the Clostridium perfringens spores.

XX Sequence 5216 BP; 1934 A; 790 C; 980 G; 1512 T; 0 other;

SQ

Query Match 15.3%; Score 38; DB 20; Length 5216;

Best Local Similarity 100.0%; Pred. No. 2.9e-08;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 38

DB 2309 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 2272

RESULT 2

AAX07549/c

ID AAX07549 standard; DNA; 5230 BP.

XX AAX07549;

XX

DT 08-JUN-1999 (first entry)

DE pTREX1 mutant beta-toxin expression cassette.

XX

XX Beta-toxin; cpb; vaccine; infection; protection;

XX retained immunogenicity; haemorrhagic enteritis; necrotic enteritis;

XX enterotoxemia; lamb dysentery; ds.

XX

OS Synthetic.

XX

XX EP892054-A1.

XX

XX 20-JAN-1999.

XX

XX 17-JUN-1998; 98EP-0202032.

XX

XX 20-JUN-1997; 97EP-0201888.

XX

XX (ALKU) AKZO NOBEL NV.

XX

XX Frandsen PL, Sergers RPAM, Waterfield NR, Wells JM;

XX

XX WPI; 1999-083571/08.

XX

XX New detoxified derivative of Clostridium perfringens beta-toxin -

XX with retained immunogenicity, useful as a vaccine to protect against

XX Clostridium perfringens infection

XX

XX Example; Fig 1c; 30pp; English.

XX

XX The sequence is that of the pTREX1 expression cassette which was used in

XX the construction of a detoxified immunogenic derivative of Clostridium

XX perfringens beta-toxin. Such a derivative is useful as a vaccine

XX to induce an immune response and protect against Clostridium

XX perfringens infection in man, pig, lamb, sheep, goat, calf and

XX bird, which causes haemorrhagic enteritis, necrotic enteritis,

XX enterotoxemia and lamb dysentery. The new beta-toxin has been

XX detoxified without impairing the immunogenicity, unlike prior

XX art vaccines which used chemical and formalin detoxifying methods.

CC

CC Production of the beta-toxin in gram positive bacteria prevents

CC prior art difficulties of isolation from dangerous Clostridium

CC perfringens, and purification from the Clostridium perfringens spores.

XX Sequence 5230 BP; 1930 A; 795 C; 983 G; 1522 T; 0 other;

SQ

Query Match 15.3%; Score 38; DB 20; Length 5230;

Best Local Similarity 100.0%; Pred. No. 2.9e-08;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 38

DB 2323 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 2286

RESULT 3

AAA15041/c

ID AAA15041 standard; DNA; 5230 BP.

XX AAA15041;

XX

DT 21-AUG-2000 (first entry)

DE Nucleotide sequence of expression plasmid pTREX1.

XX

XX Intestinal mucosa; cytokine; cytokine antagonist; anti-inflammatory;

XX Gram-positive bacteria; Lactococcus lactis; gastrointestinal tract;

XX Gram-positive bacteria; Lactococcus lactis; Crohn's disease; IL10;

XX inflammatory bowel disease; chronic colitis; Crohn's disease; IL10;

XX ulcerative colitis; tumour necrosis factor; TNF; interleukin 10; ss.

XX

OS Synthetic.

XX

XX WO200023471-A2.

XX

XX 27-APR-2000.

XX

XX 06-OCT-1999; 99WO-EP07800.

XX

XX 20-OCT-1998; 98EP-0203529.

XX

XX (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.

XX

XX Steidler L, Remaut ER, Fiers W;

XX

XX WPI; 2000-422481/36.

XX

XX Treating inflammatory bowel diseases comprises administering a

XX composition comprising a cytokine-producing or cytokine

XX antagonist-producing Gram-positive bacterial strain -

XX

XX Disclosure; Fig 1B; 45pp; English.

XX

XX The present sequence represents an expression plasmid for use in the

XX course of the invention. The specification describes an administration

XX strategy for the delivery at the intestinal mucosa of cytokines or

XX cytokine antagonists, preferably of acid sensitive anti-inflammatory

XX agents such as interleukin 10 (IL10) or tumour necrosis factor (TNF).

XX The method uses a cytokine-producing or cytokine antagonist-producing

XX Gram-positive bacterial strain (such as Lactococcus lactis). The use of

XX non-colonizing bacteria expressing cytokines or cytokine antagonists

XX allows the treatment to be directed to the disease site, whilst

XX minimizing the possibility of degradation along the gastrointestinal

XX tract. The recombinant bacteria are used in the treatment of inflammatory

XX bowel diseases, especially chronic colitis, Crohn's disease or an

XX ulcerative colitis.

XX

XX Sequence 5230 BP; 1930 A; 796 C; 982 G; 1522 T; 0 other;

XX

XX

XX Query Match 15.3%; Score 38; DB 21; Length 5230;

XX Best Local Similarity 100.0%; Pred. No. 2.9e-08;

XX Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 38

```

Db 2323 GGGGTACCGCCAGCATTTTCGGAAAAAACCCAGCGCTAAG 2286
|||||
RESULT 4
AAAX07550/c
ID AAX07550 standard; DNA: 5231 BP.
XX
AC AAX07550;
XX
DT 08-JUN-1999 (first entry)
XX
DE pTREX1A mutant beta-toxin expression cassette.
XX
KW Beta-toxin; cpb; vaccine; infection; protection;
KW retained immunogenicity; haemorrhagic enteritis; necrotic enteritis;
KW enterotoxemia; lamb dysentery; ds.
XX
OS Synthetic.
XX
PN EP892054-Al.
XX
PD 20-JAN-1999.
XX
PF 17-JUN-1998; 98EP-0202032.
XX
PR 20-JUN-1997; 97EP-0201888.
XX
PA (ALKU ) AKZO NOBEL NV.
XX
PI Frandsen PL, Sergers RPAM, Waterfield NR, Wells JM;
XX
DR WPT; 1999-083571/08.
XX
PT New detoxified derivative of Clostridium perfringens beta-toxin -
PT with retained immunogenicity, useful as a vaccine to protect against
PT Clostridium perfringens infection
XX
PS Example; Fig 1b; 30pp; English.
XX
CC The sequence is that of the pTREX1A expression cassette which was used
CC in the construction of a detoxified immunogenic derivative of Clostridium
CC perfringens beta-toxin. Such a derivative is useful as a vaccine
CC to induce an immune response and protect against Clostridium
CC perfringens infection in man, pig, lamb, sheep, goat, calf and
CC bird, which causes hemorrhagic enteritis, necrotic enteritis,
CC enterotoxemia and lamb dysentery. The new beta-toxin has been
CC detoxified without impairing the immunogenicity, unlike prior
CC art vaccines which used chemical and formalin detoxifying methods.
CC Production of the beta-toxin in gram positive bacteria prevents
CC prior art difficulties of isolation from dangerous Clostridium
CC perfringens, and purification from the Clostridium perfringens spores.
XX
SQ Sequence 5231 BP; 1930 A; 793 C; 985 G; 1523 T; 0 other;

Query Match 15.3%; Score 38; DB 20; Length 5231;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTTCGGAAAAAACCCAGCGCTAAG 38
|||||
Db 2324 GGGGTACCGCCAGCATTTTCGGAAAAAACCCAGCGCTAAG 2287
|||||
RESULT 5
AAF29646/c
ID AAF29646 standard; DNA: 5497 BP.
XX
AC AAF29646;
XX
DT 10-APR-2001 (first entry)
XX
DE Plasmid pTlmtTFF1.

```

```

XX Mouse TFF1; mTFF: trefoil protein; antiinflammatory; gastrointestinal;
XX anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;
XX Crohn's disease; ulcerative colitis; plasmid pTlmtTFF1;
XX recombinant vector; ds.
XX
OS Synthetic.
XX
PN WO200102570-Al.
XX
PD 11-JAN-2001.
XX
PF 05-JUL-2000; 2000WO-EP06343.
XX
PR 05-JUL-1999; 99EP-0870143.
XX
PA (VLAA-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX
PI Hans WC, Steidler L, Remaut ER;
XX
DR WPI; 2001-138142/14.
XX
PT Recombinant Lactococcus lactis for delivering a trefoil peptide useful
PT for treating acute or chronic gastrointestinal inflammatory diseases or
PT disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's
PT disease -
XX
PS Claim 18; Fig 1c; 44pp; English.
XX
CC The present sequence is a recombinant vector which may be used to
CC generate a recombinant Lactococcus lactis capable of delivering a trefoil
CC peptide in vivo. The recombinant microorganism is useful for
CC manufacturing an agent for the delivery of a trefoil peptide to the
CC gastrointestinal tract, and for treating gastric or intestinal diseases
CC or disorders, or lesions caused by gastric or intestinal diseases or
CC disorders. The microorganism may also be used for preparing medicament to
CC be used for treating gastric and/or gastrointestinal diseases or
CC disorders, acute gastrointestinal inflammatory diseases (e.g., acute
CC colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and
CC chronic and spontaneously recurring diseases of the gastrointestinal
CC tract comprising Crohn's disease (enteritis regionalis) and ulcerative
CC colitis (colitis ulcerosa). Disease states which can be treated by the
CC method or compositions comprising the recombinant microorganism or
CC trefoil peptides include disorders of and damage to the alimentary canal,
CC including the mouth, oesophagus, stomach and large and small intestine,
CC as well as for the protection and treatment of tissues that lie outside
CC the alimentary canal.
XX
SQ Sequence 5497 BP; 1994 A; 870 C; 1051 G; 1582 T; 0 other;

Query Match 15.3%; Score 38; DB 22; Length 5497;
Best Local Similarity 100.0%; Pred. No. 2.9e-08;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTTCGGAAAAAACCCAGCGCTAAG 38
|||||
Db 2590 GGGGTACCGCCAGCATTTTCGGAAAAAACCCAGCGCTAAG 2553
|||||
RESULT 6
AAAL5043/c
ID AAAL5043 standard; DNA: 5770 BP.
XX
AC AAAL5043;
XX
DT 21-AUG-2000 (first entry)
XX
DE Nucleotide sequence of expression plasmid pTlMIL10.
XX
KW Intestinal mucosa; cytokine; cytokine antagonist; anti-inflammatory;
KW Gram-positive bacteria; Lactococcus lactis; gastrointestinal tract;
KW inflammatory bowel disease; chronic colitis; Crohn's disease; IL10;
KW ulcerative colitis; tumour necrosis factor; TNF; interleukin 10; ss.

```

Thu Feb 13 12:35:00 2003

```

XX OS Synthetic.
XX PN WO200023471-A2.
XX PD 27-APR-2000.
XX PF 06-OCT-1999; 99WO-EP07800.
XX PR 20-OCT-1998; 98EP-0203529.
XX PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Steidler L, Remaut ER, Fiers W;
XX XX WPI; 2000-422481/36.
XX PT Treating inflammatory bowel diseases comprises administering a
XX PF composition comprising a cytokine-producing or cytokine
XX PR antagonist-producing Gram-positive bacterial strain -
XX PS Disclosure; Fig 1C; 45pp; English.
XX CC The present sequence represents an expression plasmid for use in the
XX CC course of the invention. The specification describes an administration
XX CC strategy for the delivery at the intestinal mucosa of cytokines or
XX CC cytokine antagonists, preferably of acid sensitive anti-inflammatory
XX CC agents such as interleukin 10 (IL10) or tumour necrosis factor (TNF).
XX CC The method uses a cytokine-producing or cytokine antagonist-producing
XX CC Gram-positive bacterial strain (such as Lactococcus lactis). The use of
XX CC non-colonizing bacteria expressing cytokines or cytokine antagonists
XX CC allows the treatment to be directed to the disease site, whilst
XX CC minimizing the possibility of degradation along the gastrointestinal
XX CC tract. The recombinant bacteria are used in the treatment of inflammatory
XX CC bowel diseases, especially chronic colitis, Crohn's disease or an
XX CC ulcerative colitis.
XX SQ Sequence 5770 BP; 2092 A; 922 C; 1109 G; 1647 T; 0 other;
XX CC Query Match 15.3%; Score 38; DB 21; Length 5770;
XX CC Best Local Similarity 100.0%; Pred. No. 2.9e-08;
XX CC Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 38
Db 2863 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 2826
RESULT 7
AAAL5044/c
ID AAAL5044 standard; DNA; 5870 BP.
XX AC AAAL5044;
XX DT 21-AUG-2000 (first entry)
XX DE Nucleotide sequence of expression plasmid pT1R5AH.
XX KW Intestinal mucosa; cytokine; cytokine antagonist; anti-inflammatory;
XX KW Gram-positive bacteria; Lactococcus lactis; gastrointestinal tract;
XX KW inflammatory bowel disease; chronic colitis; Crohn's disease; IL10;
XX KW ulcerative colitis; tumour necrosis factor; TNF; interleukin 10; ss.
XX OS Synthetic.
XX PN WO200023471-A2.
XX PD 27-APR-2000.
XX PF 06-OCT-1999; 99WO-EP07800.
XX PR 20-OCT-1998; 98EP-0203529.
XX PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Steidler L, Remaut ER, Fiers W;
XX XX WPI; 2000-422481/36.
XX PT Treating inflammatory bowel diseases comprises administering a
XX PF composition comprising a cytokine-producing or cytokine
XX PR antagonist-producing Gram-positive bacterial strain -
XX PS Disclosure; Fig 1C; 45pp; English.
XX CC The present sequence represents an expression plasmid for use in the
XX CC course of the invention. The specification describes an administration
XX CC strategy for the delivery at the intestinal mucosa of cytokines or
XX CC cytokine antagonists, preferably of acid sensitive anti-inflammatory
XX CC agents such as interleukin 10 (IL10) or tumour necrosis factor (TNF).
XX CC The method uses a cytokine-producing or cytokine antagonist-producing
XX CC Gram-positive bacterial strain (such as Lactococcus lactis). The use of
XX CC non-colonizing bacteria expressing cytokines or cytokine antagonists
XX CC allows the treatment to be directed to the disease site, whilst
XX CC minimizing the possibility of degradation along the gastrointestinal
XX CC tract. The recombinant bacteria are used in the treatment of inflammatory
XX CC bowel diseases, especially chronic colitis, Crohn's disease or an
XX CC ulcerative colitis.
XX SQ Sequence 5770 BP; 2092 A; 922 C; 1109 G; 1647 T; 0 other;
XX CC Query Match 15.3%; Score 38; DB 21; Length 5770;
XX CC Best Local Similarity 100.0%; Pred. No. 2.9e-08;
XX CC Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 38
Db 2863 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 2826
RESULT 8
AAAL5042/c
ID AAAL5042 standard; DNA; 5906 BP.
XX AC AAAL5042;
XX DT 21-AUG-2000 (first entry)
XX DE Nucleotide sequence of expression plasmid pT1NX.
XX KW Intestinal mucosa; cytokine; cytokine antagonist; anti-inflammatory;
XX KW Gram-positive bacteria; Lactococcus lactis; gastrointestinal tract;
XX KW inflammatory bowel disease; chronic colitis; Crohn's disease; IL10;
XX KW ulcerative colitis; tumour necrosis factor; TNF; interleukin 10; ss.
XX OS Synthetic.
XX PN WO200023471-A2.
XX PD 27-APR-2000.
XX PF 06-OCT-1999; 99WO-EP07800.
XX PR 20-OCT-1998; 98EP-0203529.
XX PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Steidler L, Remaut ER, Fiers W;
XX XX WPI; 2000-422481/36.
XX PT Treating inflammatory bowel diseases comprises administering a
XX PF composition comprising a cytokine-producing or cytokine
XX PR antagonist-producing Gram-positive bacterial strain -
XX PS Disclosure; Fig 1B; 45pp; English.

```

```

PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX Steidler L, Remaut ER, Fiers W;
XX WPI; 2000-422481/36.
XX Treating inflammatory bowel diseases comprises administering a
XX PF composition comprising a cytokine-producing or cytokine
XX PR antagonist-producing Gram-positive bacterial strain -
XX PS Disclosure; Fig 1C; 45pp; English.
XX CC The present sequence represents an expression plasmid for use in the
XX CC course of the invention. The specification describes an administration
XX CC strategy for the delivery at the intestinal mucosa of cytokines or
XX CC cytokine antagonists, preferably of acid sensitive anti-inflammatory
XX CC agents such as interleukin 10 (IL10) or tumour necrosis factor (TNF).
XX CC The method uses a cytokine-producing or cytokine antagonist-producing
XX CC Gram-positive bacterial strain (such as Lactococcus lactis). The use of
XX CC non-colonizing bacteria expressing cytokines or cytokine antagonists
XX CC allows the treatment to be directed to the disease site, whilst
XX CC minimizing the possibility of degradation along the gastrointestinal
XX CC tract. The recombinant bacteria are used in the treatment of inflammatory
XX CC bowel diseases, especially chronic colitis, Crohn's disease or an
XX CC ulcerative colitis.
XX SQ Sequence 5870 BP; 2097 A; 960 C; 1142 G; 1671 T; 0 other;
XX CC Query Match 15.3%; Score 38; DB 21; Length 5870;
XX CC Best Local Similarity 100.0%; Pred. No. 2.9e-08;
XX CC Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 38
Db 2963 GGGGTACCGCCAGCATTTCCGAAAAAACCACGCTAAG 2926
RESULT 8
AAAL5042/c
ID AAAL5042 standard; DNA; 5906 BP.
XX AC AAAL5042;
XX DT 21-AUG-2000 (first entry)
XX DE Nucleotide sequence of expression plasmid pT1NX.
XX KW Intestinal mucosa; cytokine; cytokine antagonist; anti-inflammatory;
XX KW Gram-positive bacteria; Lactococcus lactis; gastrointestinal tract;
XX KW inflammatory bowel disease; chronic colitis; Crohn's disease; IL10;
XX KW ulcerative colitis; tumour necrosis factor; TNF; interleukin 10; ss.
XX OS Synthetic.
XX PN WO200023471-A2.
XX PD 27-APR-2000.
XX PF 06-OCT-1999; 99WO-EP07800.
XX PR 20-OCT-1998; 98EP-0203529.
XX PA (VLAAs-) VLAAMS INTERUNIVERSITAIR INST BIOTECHNOG.
XX PI Steidler L, Remaut ER, Fiers W;
XX XX WPI; 2000-422481/36.
XX PT Treating inflammatory bowel diseases comprises administering a
XX PF composition comprising a cytokine-producing or cytokine
XX PR antagonist-producing Gram-positive bacterial strain -
XX PS Disclosure; Fig 1B; 45pp; English.

```

The present sequence represents an expression plasmid for use in the course of the invention. The specification describes an administration strategy for the delivery at the intestinal mucosa of cytokines or cytokine antagonists, preferably of acid sensitive anti-inflammatory agents such as interleukin 10 (Il10) or tumour necrosis factor (TNF). The method uses a cytokine-producing or cytokine antagonist-producing Gram-positive bacterial strain (such as *Lactococcus lactis*). The use of non-colonizing bacteria expressing cytokines or cytokine antagonists allows the treatment to be directed to the disease site, whilst minimizing the possibility of degradation along the gastrointestinal tract. The recombinant bacteria are used in the treatment of inflammatory bowel diseases, especially chronic colitis, Crohn's disease or an ulcerative colitis.

Query Match 15.3%; Score 38; DB 21; Length 5906;
 Best Local Similarity 100.0%; Pred. No. 2.9e-08;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTACCGCCAGCATTTTCGAAAAAACCACGCTAAG 38
 Db 2999 GGGTACCGCCAGCATTTTCGAAAAAACCACGCTAAG 2962

RESULT 9
 AAX13472
 ID AAX13472 standard; DNA; 1566 BP.
 AC AAX13472;
 XX
 XX
 DT 19-MAR-1999 (first entry)
 DE Enterococcus faecalis genome contig SEQ ID NO:535.
 DE Enterococcus faecalis; contig; detection; Enterococcal infection;
 KW vaccine; attenuation; computer readable medium; ds.
 XX
 OS Enterococcus faecalis.
 XX
 PN WO9850555-A2.
 XX
 PD 12-NOV-1998.
 XX
 PF 04-MAY-1998; 98WO-US08985.
 XX
 PR 14-NOV-1997; 97US-0066009.
 PR 06-MAY-1997; 97US-0044031.
 PR 16-MAY-1997; 97US-0046655.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Barash SC, Billon PJ, Kunsch CA;
 XX
 DR WPI; 1999-045171/04.
 XX
 PT New isolated Enterococcus faecalis polynucleotides and polypeptides
 PT - used to develop products for the detection of Enterococcus and for
 PT use in vaccines for prevention or attenuation of Enterococcus
 PT infection.
 XX
 PS Claim 1; Page 1787-1788; 2084pp; English.
 XX
 CC A computer readable medium has been developed which has recorded on it
 CC 982 nucleotide sequences isolated from the Enterococcus faecalis genome.
 CC AAX12938 to AAX13919 represent these nucleotide sequences which are
 CC primary nucleotide sequences, also known as contigs. The computer-based
 CC system can identify fragments of the Enterococcus faecalis genome with
 CC commercial importance. The products can be used to detect the presence
 CC of Enterococcus faecalis in samples. They can also be used for
 CC diagnosing Enterococcal infection in an animal and monitoring
 CC progression of disease, and for identifying agents which can be used to

CC modulate the growth or pathogenicity of Enterococcus faecalis, or
 CC another related organism, in vivo or in vitro. In particular the
 CC polypeptides encoded by the Enterococcus faecalis nucleotide sequences
 CC can be used in vaccines to prevent or attenuate an Enterococcal
 CC infection.
 XX
 SQ Sequence 1566 BP; 460 A; 273 C; 260 G; 559 T; 14 other;
 Query Match 7.6%; Score 19; DB 20; Length 1566;
 Best Local Similarity 100.0%; Pred. No. 21;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 TTTTATACAAAAATAAT 116
 Db 1489 TTTTATACAAAAATAAT 1507

RESULT 10
 AAS46664/c
 ID AAS46664 standard; DNA; 6076 BP.
 XX
 AC AAS46664;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Tumour suppressor gene derived chemically modified sequence #386.
 XX
 KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200168912-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 15-MAR-2001; 2001WO-EP02955.
 XX
 PR 15-MAR-2000; 2000DE-1013847.
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2001-602752/68.
 XX
 PT Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer -
 XX
 PS Claim 1; SEQ ID No 386; 27pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (Cp DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (SS) and sequences complementary to (ss). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the

CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6076 BP; 2013 A; 73 C; 1223 G; 2767 T; 0 other;

Query Match 7.6%; Score 19; DB 22; Length 6076;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AAAATAATACACACAATA 127
|||||
DB 3018 AAAATAATACACACAATA 3000

RESULT 11
ABK40042/c
ID ABK40042 standard; DNA; 6136 BP.
XX
AC ABK40042;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #62 strand 2.
XX
KW Human; ds; bisulphite treatment; CpG; DNA methylation; cancer; tumour;
KW cytostatic; ALDH6; CYP11A; CYP11B1; CYP3A3; DPYD; EPHX2; OCLN; TXNRD1;
KW UGT8; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07470.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154757/20.
XX
PT New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PT useful for detecting cytosine methylation state of genes associated
PT with pharmacogenomics and for therapy of diseases e.g. cancer .
XX
PS Claim 1; SEQ ID No 124; 24pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DPYD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996), and their complementary sequences, or a sequence (S1) chosen
CC from 87 sequences and their complements. The chemical pretreatment
CC is bisulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides

CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements, arranged in
CC an array for analysing diseases associated with the methylation state
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 nucleic acids. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 6136 BP; 1775 A; 69 C; 1285 G; 3007 T; 0 other;

Query Match 7.6%; Score 19; DB 24; Length 6136;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 ATACAAAAATAATACAAAC 121
|||||
DB 530 ATACAAAAATAATACAAAC 512

RESULT 12
AAS46531
ID AAS46531 standard; DNA; 7008 BP.
XX
AC AAS46531;
XX
DT 18-DEC-2001 (first entry)
XX
DE Tumour suppressor gene derived chemically modified sequence #253.
XX
KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic;
KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
KW cytosine methylation; ds.
XX
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
XX
PR 06-APR-2000; 2000DE-1019038.
XX
PR 07-APR-2000; 2000DE-1019173.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer .
XX
PS Claim 1; SEQ ID No 253; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (SS) and sequences complementary to (SS). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state

CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 7008 BP; 2322 A; 26 C; 1306 G; 3354 T; 0 other;
 Query Match 7.6%; Score 19; DB 22; Length 7008;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 142 TTAGGTGTTATAATAATA 160
 |||||
 Db 2443 TTAGGTGTTATAATAATA 2461

RESULT 13
 AAS98600/C
 ID AAS98600 standard; DNA; 140036 BP.

XX AC AAS98600;

DT 12-MAR-2002 (first entry)

XX DE Human genomic DNA for PHIP/NDRP.

XX KW PHIP: Pleckstrin homology domain-interacting protein; NDRP: ds;
 KW neuronal differentiation-related protein; insulin receptor substrate;
 KW IRS: signal transducer and activator of transcription; STAT;
 KW transgenic animal; diabetes mellitus type 2; hyperglycaemia;
 KW myotonic muscular dystrophy; acanthosis; nigricans; retinopathy;
 KW nephropathy; artherosclerosis; peripheral arterial disease; cancer;
 KW adenocarcinoma; leukaemia; breast cancer; prostate cancer; colon cancer;
 KW ovarian cancer; autoimmune disease; inflammation; immunodeficiency.

XX OS Homo sapiens.

XX PN WO200185785-A2.

XX PD 15-NOV-2001.

PF 10-MAY-2001; 2001WO-CA00673.

XX PR 11-MAY-2000; 2000US-203561P.

XX PA (ROZA/) ROZAKIS-ADCOCK M.

XX PA (FARH/) FARHANG-FALLAH J.

XX PA (CHEN/) CHENG A.

XX PI Rozakis-Adcock M, Farhang-Fallah J, Cheng A;

XX DR WPI; 2002-041586/05.

XX PT Novel Pleckstrin homology domain interacting protein recruiting
 PT proteins of insulin receptor substrate family, and signal transducer
 PT and activator of transcription factors to their receptors, useful to
 PT treat diabetes

XX PS Disclosure; Page 99-133; 139pp; English.

XX CC The invention relates to an isolated pleckstrin homology domain
 CC interacting protein (PHIP) that recruits proteins of the insulin

CC receptor substrate (IRS) family, and signal transducer and activator of
 CC transcription (STAT) transcription factors, to receptors that interact
 CC with and phosphorylate the proteins and STAT transcription factors,
 CC the nucleic acid encoding PHIP (nPHIP), a nucleic acid which binds to
 CC nPHIP or regions of it, analogues, fragments or allelic variants of PHIP
 CC or nPHIP, a nucleic acid sequence having substantial sequence identity or
 CC sequence similarity with a nucleic acid sequence fully defined
 CC human neuronal differentiation-related protein (NDRP) nucleic acid
 CC sequence or its exons as given in the specification, expression
 CC vectors and host cells expressing the nucleic acids, anti-PHP antibodies,
 CC and a transgenic animal not already expressing PHIP. The nucleic
 CC acids, proteins and antibodies are useful for diagnosis and treatment of
 CC a condition associated with an insulin receptor (e.g. diabetes mellitus
 CC type 2, hyperglycaemia, myotonic muscular dystrophy, acanthosis,
 CC nigricans, retinopathy, nephropathy, artherosclerosis, peripheral
 CC arterial disease) or cancer (e.g. adenocarcinoma, leukaemia, breast
 CC cancer, prostate cancer, colon cancer, ovarian cancer and many others
 CC given in the specification), autoimmune disease, inflammation and
 CC immunodeficiency. The protein is also useful for discovering
 CC or testing compounds which may be either enhancers or inhibitors of PHIP
 CC function. The present sequence is genomic DNA encoding PHIP and NDRP.

XX SQ Sequence 140036 BP; 48500 A; 25843 C; 22499 G; 43194 T; 0 other;

Query Match 7.6%; Score 19; DB 24; Length 140036;
 Best Local Similarity 100.0%; Pred No. 17;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 47 AGTTAAATAATCAGAAAT 65
 |||||

Db 128998 AGTTAAATAATCAGAAAT 128980

RESULT 14

AAC22654

ID AAC22654 standard; cDNA; 251 BP.

XX AC AAC22654;

DT 06-OCT-2000 (first entry)

DE Human secreted protein 5' EST, SEQ ID NO: 26729.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping; ss.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI; 2000-500381/45.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 1; SEQ ID 26729; 71pp + CD-ROM; English.

XX CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)

CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors.
XX
SQ Sequence 251 BP; 92 A; 38 C; 52 G; 69 T; 0 other;

Query Match 7.2%; Score 18; DB 21; Length 251;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 TTGTATTAGGTGTATAA 154
|||||
Db 81 TTGTATTAGGTGTATAA 98

RESULT 15
ABN16703
ID ABN16703 standard; cDNA; 286 BP.

AC ABN16703;

XX 24-JUN-2002 (first entry)

DE Human ORFX polynucleotide sequence SEQ ID NO:1883.

XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
KW hypertension; hypothyroidism; cholesterol ester storage disease;
KW immune deficiency; immune disorder; infectious disease;
KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
KW myasthenia gravis; gene; ss.

XX Homo sapiens.

XX WO200192523-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US10836.

XX 30-MAY-2000; 2000US-206132P.

XX 29-AUG-2000; 2000US-228716P.

XX (CURA-) CURAGEN CORP.

XX Shimkets RA, Leach MD;

XX WPI; 2002-106308/14.

XX P-PSDB; ABP00951.

XX Novel human polypeptides and polynucleotides useful for diagnosing,
XX preventing and treating cardiovascular disease, neurodegenerative,
XX hyperproliferative disorders and autoimmune disorders -

XX Disclosure; SEQ ID 1883; 1037pp; English.

XX The present invention describes substantially purified human proteins
XX (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
XX in the specification). ABN15762 to ABN27252 encode the human ORFX
XX proteins given in ABP00010 to ABP11500. ORFX proteins are useful for
XX treating or preventing a pathology associated with an ORFX-associated
XX disorder in humans, and in the manufacture of a medicament for treating a
XX syndrome associated with ORFX-associated disorder. ORFX polynucleotide
XX sequences can be used in gene therapy. ORFX sequences can be used in the
XX treatment of cancer, hyperproliferative disorders, cirrhosis of liver,

CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
CC osteoarthritis, neurodegenerative disorders, disorders related to organ
CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
CC storage disease, various immune deficiencies and disorders, infectious
CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
CC bone degenerative disorders, or periodontal disease, and for gut
CC protection or regeneration and treatment of lung or liver fibrosis.
CC reperfusion injury in various tissues and conditions resulting from
CC systemic cytokine damage.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 286 BP; 77 A; 71 C; 58 G; 80 T; 0 other;

Query Match 7.2%; Score 18; DB 24; Length 286;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 TTTTATTATACAAAAATA 114

|||||

Db 140 TTTTATTATACAAAAATA 157

Search completed: February 13, 2003, 04:59:25
Job time : 340 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.
OM nucleic - nucleic search, using sw model
Run on: February 13, 2003, 04:52:47 ; Search time 67 seconds
(without alignments)
1139.739 Million cell updates/sec
Title: US-09-895-435-4
Perfect score: 249
Sequence: 1 ggggtaccgcagcatttcg.....gacgcggaggaaatcacatg 249
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0
Searched: 441362 seqs, 153338381 residues
Word size: 0

Total number of hits satisfying chosen parameters: 882724
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	17	6.8	439	4	US-09-397-787-269
C 2	17	6.8	1029	4	US-09-330-611-3
C 3	17	6.8	12124	1	US-08-181-271A-36
C 4	17	6.8	12124	1	US-08-449-315-36
C 5	17	6.8	12124	1	US-08-444-803-36
C 6	17	6.8	12124	1	US-08-449-043-36
C 7	17	6.8	12124	1	US-08-456-265A-36
C 8	17	6.8	12124	1	US-08-455-416-36
C 9	17	6.8	12124	1	US-08-455-244-36
C 10	17	6.8	12124	1	US-08-454-876-36
C 11	17	6.8	12124	2	US-08-457-364-36
C 12	17	6.8	12124	2	US-08-456-262-36
C 13	17	6.8	12124	2	US-08-456-240-36
C 14	17	6.8	12124	2	US-08-455-736-36
C 15	17	6.8	12124	2	US-08-971-217-36
C 16	17	6.8	12124	4	US-09-350-600-36
C 17	17	6.8	18596	4	US-09-318-448-11
C 18	17	6.8	50000	4	US-09-146-053-4
C 19	17	6.8	168575	4	US-09-426-290-1
C 20	16	6.4	240	4	US-09-134-001C-2403
C 21	16	6.4	264	4	US-09-134-001C-2149
C 22	16	6.4	319	1	US-07-593-637-14
C 23	16	6.4	417	4	US-09-134-001C-482
C 24	16	6.4	978	4	US-08-858-207A-174
C 25	16	6.4	999	4	US-08-961-527-193
C 26	16	6.4	1241	1	US-07-593-657-6
C 27	16	6.4	1241	4	US-08-942-012B-3

28 16 6.4 2061 4 US-09-653-839-7 Sequence 7, Appl
29 16 6.4 2109 4 US-09-653-839-5 Sequence 5, Appl
30 16 6.4 2172 4 US-09-653-839-3 Sequence 3, Appl
31 16 6.4 2220 4 US-09-653-839-1 Sequence 1, Appl
32 16 6.4 2510 4 US-08-894-324A-2 Sequence 2, Appl
33 16 6.4 2806 4 US-09-653-839-9 Sequence 9, Appl
34 16 6.4 3839 4 US-09-245-248B-57 Sequence 57, Appl
C 35 16 6.4 3854 2 US-08-720-484A-1 Sequence 1, Appl
C 36 16 6.4 3854 3 US-08-953-823A-1 Sequence 1, Appl
C 37 16 6.4 3854 4 US-09-398-239-1 Sequence 1, Appl
38 16 6.4 4337 3 US-09-187-049-1 Sequence 1, Appl
39 16 6.4 5181 1 US-08-257-073-10 Sequence 10, Appl
40 16 6.4 13965 4 US-09-453-702B-48 Sequence 48, Appl
41 16 6.4 15225 2 US-08-892-403A-2 Sequence 2, Appl
42 16 6.4 99500 4 US-09-798-096-10 Sequence 10, Appl
43 15 6.0 339 4 US-09-134-001C-1040 Sequence 1040, Ap
C 44 15 6.0 432 4 US-09-071-035-129 Sequence 129, App
45 15 6.0 453 4 US-09-134-001C-2262 Sequence 2262, Ap

ALIGNMENTS

RESULT 1
US-09-397-787-269/c
; Sequence 269, Application US/09397787
; Patent No. 6468758
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: King, Gordon E.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR OVARIAN
; FILE REFERENCE: 210121.466C2
; CURRENT APPLICATION NUMBER: US/09/397,787
; CURRENT FILING DATE: 1999-09-16
; NUMBER OF SEQ ID NOS: 334
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 269
; LENGTH: 439
; TYPE: DNA
; ORGANISM: Homo sapien
; NAME/KEY: misc_feature
; LOCATION: (1)...(439)
; OTHER INFORMATION: n = A,T,C or G
US-09-397-787-269

Query Match 6.8%; Score 17; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 ATTCTTGCACATACA 90
Db 142 ATTCTTGCACATACA 126
|||||

RESULT 2
US-09-330-611-3
; Sequence 3, Application US/09330611
; Patent No. 6248874
; GENERAL INFORMATION:
; APPLICANT: FREY, Perry A.
; APPLICANT: RUZICKA, Frank J.
; TITLE OF INVENTION: DNA MOLECULES ENCODING BACTERIAL LYSINE 2,3-AMINOMUTASE
; FILE REFERENCE: 032026/0476
; CURRENT APPLICATION NUMBER: US/09/330,611
; CURRENT FILING DATE: 1999-06-11
; EARLIER APPLICATION NUMBER: US 09/198,942
; EARLIER FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 3
; LENGTH: 1029
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1026)
US-09-330-611-3

Query Match          6.8%; Score 17; DB 4; Length 1029;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 226 TATAGACGGGAGGAAA 242
Db 102 TATAGACGGGAGGAAA 118

RESULT 3
US-08-181-271A-36/c
; Sequence 36, Application US/08181271A
; Patent No. 5614395
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,271A
; FILING DATE: 13-JAN-94
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667

; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cucumis sativus
; INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: pBscuccchrt5
US-08-181-271A-36

Query Match          6.8%; Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 AAAAAATCAGAAATAT 67
Db 11018 AAAAAATCAGAAATAT 11002

RESULT 4
US-08-449-315-36/c
; Sequence 36, Application US/08449315
; Patent No. 5650505
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
```

APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,315
FILING DATE: 24-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBsgucchrht5
US-08-449-315-36
Query Match 6.8%; Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 51 AAAAAATCAGAAAATAT 67
|||||
Db 11018 AAAAAATCAGAAAATAT 11002
RESULT 5
US-08-444-803-36/c
Sequence 36, Application US/08444803
Patent No. 5654414
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,803
FILING DATE: 19-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993

us-09-895-435-4.rni

Thu Feb 13 12:35:00 2003

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cucumis sativus
; INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: pBSCucchrht5
; US-08-444-803-36

Query Match 6.8%; Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AAAAAATCAGAAATAT 67
Db 11018 AAAAAATCAGAAATAT 11002

RESULT 6
US-08-449-043-36/C
; Sequence 36, Application US/08449043
; Patent No. 5689044

```

```

; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,043
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-94
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:

```

APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 35,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBScucchrcht5
US-08-449-043-36

Query Match 6.8%; Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AAAAAATCAGAAATAT 67
|||||

Db 11018 AAAAAATCAGAAATAT 11002

RESULT 7
US-08-456-265A-36/c
Sequence 36, Application US/08456265A
Patent No. 5767369
GENERAL INFORMATION:
APPLICANT: Alexander, Danny C.
APPLICANT: Ryals, John A.
APPLICANT: Goodman, Robert M.
APPLICANT: Stinson, Jeffrey R.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 520 White Plains Road, P.O. Box 2005
CITY: Tarrytown
STATE: New York
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/456,265A
FILING DATE: 31-MAY-95
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/181,271
FILING DATE: 13-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727/DIV10
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBScucchrcht5
US-08-456-265A-36

Query Match 6.8%; Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 51 AAAAAATCAGAAATAT 67
|||||

Db 11018 AAAAAATCAGAAATAT 11002

RESULT 8
US-08-455-416-36/c
Sequence 36, Application US/08455416
Patent No. 577200
GENERAL INFORMATION:

Thu Feb 13 12:35:00 2003

APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,416
FILING DATE: 31-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672

FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBScucchrcht5
US-08-455-416-36
Query Match 6.8% Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 51 AAAAAATCAGAAATAT 67
|||||
Db 11018 AAAAAATCAGAAATAT 11002
RESULT 9
US-08-455-244-36/c
; Sequence 36, Application US/08455244
; Patent No. 5789214
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/455,244
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBScucchrht5
US-08-455-244-36

Query Match 6.8%; Score 17; DB 1; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 AAAAAATCAGAAATAT 67
|||||
Db 11018 AAAAAATCAGAAATAT 11002
RESULT 10
US-08-454-876-36/c
Sequence 36, Application US/08454876
Patent No. 5804593
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/454,876
FILING DATE: 31-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989

Thu Feb 13 12:35:00 2003

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cucumis sativus
; INDIVIDUAL ISOLATE: Cucurbit Chitinase Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: pBScuccchrht5
;
; US-08-454-876-36
;
; Query Match 6.8%; Score 17; DB 1; Length 12124;
; Best Local Similarity 100.0%; Pred. No. 16;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 51 AAAAAATCAGAAATAT 67
; |||||
; DB 11018 AAAAAATCAGAAATAT 11002
;
; RESULT 11
; US-08-457-364-36/C
; Sequence 36, Application US/08457364
; Patent No. 5847258
; GENERAL INFORMATION:
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,364
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-94
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,157
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```



```
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cucumis sativus
; INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: pBScucchrht5
US-08-457-364-36
```

```
Query Match 6.8%; Score 17; DB 2; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 51 AAAAAATCAGAAATAT 67
      |||||||
Db 11018 AAAAAATCAGAAATAT 11002
```

RESULT 12

```
US-08-456-262-36/c
; Sequence 36, Application US/08456262
; Patent No. 5851766
```

GENERAL INFORMATION:

```
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
```

COMPUTER READABLE FORM:

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/456,262
; FILING DATE: 31-MAY-1995
```

CLASSIFICATION: 435

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-94
; APPLICATION NUMBER: US 08/093,301
; FILING DATE: 16-JUL-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/937,197
; FILING DATE: 6-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/678,378
; FILING DATE: 1-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/305,566
; FILING DATE: 6-FEB-1989
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: US 07/165,667
; FILING DATE: 8-MAR-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/042,847
; FILING DATE: 6-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/632,441
; FILING DATE: 21-DEC-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/425,504
; FILING DATE: 20-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/848,506
; FILING DATE: 6-MAR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/768,122
; FILING DATE: 27-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/580,431
; FILING DATE: 7-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/368,672
; FILING DATE: 20-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/329,018
; FILING DATE: 24-MAR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/045,957
; FILING DATE: 12-APR-1993
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cucumis sativus
; INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: pBScucchrht5
US-08-456-262-36
```

```
Query Match 6.8%; Score 17; DB 2; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 51 AAAAAATCAGAAATAT 67
      |||||||
```

```
Db 11018 AAAAAATCAGAAATAT 11002
```

RESULT 13

```
US-08-456-240-36/c
; Sequence 36, Application US/08456240
; Patent No. 5856154
```

GENERAL INFORMATION:

```
; APPLICANT: Ryals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
```

```

;
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Cucumis sativus
; INDIVIDUAL ISOLATE: Cucumer Chitinase Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: pBscuccchrht5
;
US-08-456-240-36
Query Match 6.8%; Score 17; DB 2; Length 12124;
Best Local Similarity 100.0%; Prod.No.16;
Matches 17; Conservative 0; Mismatches 0; Indels 0;

QY 51 AAAAAATCAGAAATAT 67
|||||
Db 11018 AAAAAATCAGAAATAT 11002

RESULT 14
US-08-455-736-36/c
; Sequence 36, Application US/08455736
; Patent No. 5880328
; GENERAL INFORMATION:
; APPLICANT: Rvals, John A.
; APPLICANT: Alexander, Danny C.
; APPLICANT: Beck, James J.
; APPLICANT: Duesing, John H.
; APPLICANT: Friedrich, Leslie B.
; APPLICANT: Goodman, Robert M.
; APPLICANT: Harms, Christian
; APPLICANT: Meins, Jr., Frederick
; APPLICANT: Montoya, Alice
; APPLICANT: Moyer, Mary B.
; APPLICANT: Neuhaus, Jean-Marc
; APPLICANT: Payne, George B.
; APPLICANT: Sperison, Christoph
; APPLICANT: Stinson, Jeffrey R.
; APPLICANT: Uknes, Scott J.
; APPLICANT: Ward, Eric R.
; APPLICANT: Williams, Shericca C.
; TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
; TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,736
; FILING DATE: 31-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/181,271
; FILING DATE: 13-JAN-1994
; APPLICATION NUMBER: US 08/093,301

```

FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/PL/COC 1727
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBSccuchrht5
US-08-455-736-36

Query Match 6.8% Score 17; DB 2; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 AAAAAATCAGAAATAT 67
|||||
Db 11018 AAAAAATCAGAAATAT 11002

RESULT 15
US-08-971-217-36/c
; Sequence 36, Application US/08971217

Patent No. 5942662
GENERAL INFORMATION:
APPLICANT: Rvals, John A.
APPLICANT: Harms, Christian
APPLICANT: Friedrich, Leslie
APPLICANT: Beck, James
APPLICANT: Uknes, Scott
APPLICANT: Ward, Eric
TITLE OF INVENTION: INDUCIBLE HERBICIDE RESISTANCE
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: NO. 5942662artis Corporation
STREET: 3054 Cornwallis Road, P.O. Box 12257
CITY: Research Triangle Park
STATE: NC
COUNTRY: USA
ZIP: 27709
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/971,217
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,364
FILING DATE: 31-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/181,271
FILING DATE: 13-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957

Thu Feb 13 12:35:00 2003

FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727/DIV5/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8587
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 12124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cucumis sativus
INDIVIDUAL ISOLATE: Cucumber Chitinase Genomic DNA
IMMEDIATE SOURCE:
CLONE: pBscucchrt5
US-08-971-217-36

Query Match 6.8%; Score 17; DB 2; Length 12124;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 51 AAAAAATCAGAAATAT 67
|||||
Db 11018 AAAAAATCAGAAATAT 11002

Search completed: February 13, 2003, 06:42:21
Job time : 89 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2003, 04:56:57 ; Search time 114 Seconds
(without alignments)
1112.470 Million cell updates/sec

Title: US-09-895-435-4
Perfect score: 249
Sequence: 1 999gtaccgcagcatttcg.....gacgcggagggaatcacatg 249

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 424239 seqs, 254661826 residues

Word size : 0

Total number of hits satisfying chosen parameters: 848478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications, NA:*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	38	15.3	5230	10	US-09-838-718A-5
C 2	38	15.3	5770	10	US-09-838-718A-7
C 3	38	15.3	5870	10	US-09-838-718A-8
C 4	38	15.3	5906	10	US-09-838-718A-6
C 5	19	7.6	1566	10	US-09-070-927A-535
C 6	18	7.2	341	10	US-09-960-352-2720
C 7	18	7.2	457	10	US-09-864-761-3543
C 8	18	7.2	1500	9	US-09-938-842A-4541
C 9	18	7.2	2000	9	US-09-938-842A-4002
C 10	18	7.2	4704	10	US-09-764-877-3292
C 11	18	7.2	116592	10	US-09-818-512-3
C 12	18	7.2	640681	10	US-09-790-988-1
C 13	17	6.8	148	10	US-09-962-436-517
C 14	17	6.8	309	10	US-09-969-373-1246
C 15	17	6.8	433	10	US-09-974-300-5415
C 16	17	6.8	439	10	US-09-876-889-269
C 17	17	6.8	1029	9	US-09-847-010-3
C 18	17	6.8	2000	9	US-09-938-842A-2876
C 19	17	6.8	2000	9	US-09-938-842A-3160

C 20	17	6.8	2389	9	US-10-098-841-192
C 21	17	6.8	2531	10	US-09-070-927A-96
C 22	17	6.8	4610	9	US-09-984-001-1
C 23	17	6.8	4735	10	US-09-967-552A-71
C 24	17	6.8	5676	10	US-09-764-877-3756
C 25	17	6.8	6019	10	US-09-764-847-1719
C 26	17	6.8	6479	9	US-10-098-841-101
C 27	17	6.8	11057	9	US-09-764-868-1308
C 28	17	6.8	18596	9	US-09-954-531-124
C 29	17	6.8	18596	9	US-09-954-531-348
C 30	17	6.8	18596	10	US-09-880-107-1590
C 31	17	6.8	18596	10	US-09-967-768A-119
C 32	17	6.8	41104	10	US-09-816-685-3
C 33	17	6.8	99916	10	US-09-816-685-3
C 34	17	6.8	180557	12	US-10-003-806-6
C 35	17	6.8	180557	12	US-10-003-806-6
C 36	17	6.8	197957	10	US-09-822-246-3
C 37	16	6.4	22	9	US-10-113-877-174
C 38	16	6.4	222	10	US-09-974-300-4140
C 39	16	6.4	250	10	US-09-783-590-16
C 40	16	6.4	257	10	US-09-964-824A-69
C 41	16	6.4	301	10	US-09-960-352-13603
C 42	16	6.4	328	10	US-09-967-768A-302
C 43	16	6.4	368	10	US-09-867-701-3705
C 44	16	6.4	381	9	US-09-796-692-7000
C 45	16	6.4	391	10	US-09-983-965-503

ALIGNMENTS

RESULT 1

US-09-838-718A-5/c

; Sequence 5, Application US/09838718A

; Patent No. US20020019043A1

; GENERAL INFORMATION:

; APPLICANT: Steidler, Lothar

; APPLICANT: Remaut, Erik R.

; APPLICANT: Fiers, Walter R.

; TITLE OF INVENTION: USE OF A CYTOKINE-PRODUCING LACTOCOCCUS STRAIN TO TREAT COLITIS

; FILE REFERENCE: 2676-4779US

; CURRENT APPLICATION NUMBER: US/09/838,718A

; PRIOR FILING DATE: 2001-04-19

; PRIOR APPLICATION NUMBER: PCT/EP99/07800

; PRIOR FILING DATE: 1999-10-06

; PRIOR APPLICATION NUMBER: EP 98203529.7

; PRIOR FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 8

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 5

; LENGTH: 5230

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTREX1

US-09-838-718A-5

Query Match 15.3%; Score 38; DB 10; Length 5230;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGTACCGCCAGCATTTCGGAAAAAACACCGCTAAG 38
|||||

Db 2323 GGGGTACCGCCAGCATTTCGGAAAAAACACCGCTAAG 2286
|||||

RESULT 2

US-09-838-718A-7/c

; Sequence 7, Application US/09838718A

; Patent No. US20020019043A1

; GENERAL INFORMATION:

; APPLICANT: Steidler, Lothar

; APPLICANT: Remaut, Erik R.

Thu Feb 13 12:35:01 2003

```

; APPLICANT: Fiers, Walter R.
; TITLE OF INVENTION: USE OF A CYTOKINE-PRODUCING LACTOCOCCUS STRAIN TO TREAT COLITIS
; FILE REFERENCE: 2676-4779US
; CURRENT APPLICATION NUMBER: US/09/838,718A
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: PCT/EP99/07800
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: EP 98203529.7
; PRIOR FILING DATE: 1998-10-20
; PRIOR APPLICATION NUMBER: EP 98203529.7
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; QUERY ID NO 7
; BEST LOCAL SIMILARITY 100.0%; Pred. No. 2.5e-09;
; Mismatches 0; Conservative 0; Indels 0; Gaps 0;
; LENGTH: 5770
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTLM1L10
; US-09-838-718A-7

Query Match 15.3%; Score 38; DB 10; Length 5770;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTACCGCCAGCATTTCCGAAAAAAACACCGCTAAG 38
Db 2863 GGGTACCGCCAGCATTTCCGAAAAAAACACCGCTAAG 2826

RESULT 3
US-09-838-718A-8/c
; Sequence 8, Application US/09838718A
; Patent No. US20020019043A1
; GENERAL INFORMATION:
; APPLICANT: Steidler, Lothar
; APPLICANT: Remaut, Erik R.
; APPLICANT: Fiers, Walter R.
; TITLE OF INVENTION: USE OF A CYTOKINE-PRODUCING LACTOCOCCUS STRAIN TO TREAT COLITIS
; FILE REFERENCE: 2676-4779US
; CURRENT APPLICATION NUMBER: US/09/838,718A
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: PCT/EP99/07800
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: EP 98203529.7
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; QUERY ID NO 8
; BEST LOCAL SIMILARITY 100.0%; Pred. No. 2.5e-09;
; Mismatches 0; Conservative 0; Indels 0; Gaps 0;
; LENGTH: 5870
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTITR5AH
; US-09-838-718A-8

Query Match 15.3%; Score 38; DB 10; Length 5870;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTACCGCCAGCATTTCCGAAAAAAACACCGCTAAG 38
Db 2963 GGGTACCGCCAGCATTTCCGAAAAAAACACCGCTAAG 2926

RESULT 4
US-09-838-718A-6/c
; Sequence 6, Application US/09838718A
; Patent No. US20020019043A1
; GENERAL INFORMATION:
; APPLICANT: Steidler, Lothar
; APPLICANT: Remaut, Erik R.
; APPLICANT: Fiers, Walter R.
; TITLE OF INVENTION: USE OF A CYTOKINE-PRODUCING LACTOCOCCUS STRAIN TO TREAT COLITIS
; FILE REFERENCE: 2676-4779US

```

```

; CURRENT APPLICATION NUMBER: US/09/838,718A
; CURRENT FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: PCT/EP99/07800
; PRIOR FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: EP 98203529.7
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; QUERY ID NO 6
; BEST LOCAL SIMILARITY 100.0%; Pred. No. 2.5e-09;
; Mismatches 0; Conservative 0; Indels 0; Gaps 0;
; LENGTH: 5906
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid pTINX
; US-09-838-718A-6

Query Match 15.3%; Score 38; DB 10; Length 5906;
Best Local Similarity 100.0%; Pred. No. 2.5e-09;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGTACCGCCAGCATTTCCGAAAAAAACACCGCTAAG 38
Db 2999 GGGTACCGCCAGCATTTCCGAAAAAAACACCGCTAAG 2962

RESULT 5
US-09-070-927A-535
; Sequence 535, Application US/09070927A
; Patent No. US20020120116A1
; GENERAL INFORMATION:
; APPLICANT: Charles A. Kunsch
; APPLICANT: Patrick J. Dillon
; APPLICANT: Steven Barash
; TITLE OF INVENTION: Enterococcus faecialis Polynucleotides and Polypeptides
; NUMBER OF SEQUENCES: 982
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,927A
; FILING DATE: 04-May-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/046,655
; FILING DATE: 1997-05-16
; APPLICATION NUMBER: 60/044,031
; FILING DATE: 1997-05-06
; APPLICATION NUMBER: 60/066,009
; FILING DATE: 1997-11-14
; ATTORNEY/AGENT INFORMATION:
; NAME: Kenley K. Hoover
; REGISTRATION NUMBER: 40,302
; REFERENCE/DOCKET NUMBER: PB369
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 535:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1566 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 535:
; US-09-070-927A-535

```

```

Query Match          7.6%; Score 19; DB 10; Length 1566;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 TTTTATACAAAAATAAT 116
   |||||||
DB 1489 TTTTATACAAAAATAAT 1507

RESULT 6
US-09-960-352-2720/c
; Sequence 2720, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 2720
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 12-LIB34-013-Q1-EL-C7
US-09-960-352-2720

Query Match          7.2%; Score 18; DB 10; Length 341;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 TTAGACGCAACACATTT 226
   |||||||
DB 253 TTAGACGCAACACATTT 236

RESULT 7
US-09-864-761-3543
; Sequence 3543, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3543
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007966.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.61
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.71
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.91
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.99
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
US-09-864-761-3543

Query Match          7.2%; Score 18; DB 10; Length 457;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 TTGTATTAGGTGTATAA 154
   |||||||
DB 15 TTGTATTAGGTGTATAA 32

RESULT 8
US-09-938-842A-4541
; Sequence 4541, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4541
; LENGTH: 1500
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4541

```

Query Match 7.2%; Score 18; DB 9; Length 1500;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 104 TACAAAAAATAATACAC 121
 |||||
 Db 419 TACAAAAAATAATACAC 436

RESULT 9
 US-09-938-842A-4002
 ; Sequence 4002, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIPL300-3
 ; CURRENT APPLICATION NUMBER: US/09/938.842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 4002
 ; LENGTH: 2000
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-4002

Query Match 7.2%; Score 18; DB 9; Length 2000;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 TTTTATACAAAAATA 114
 |||||
 Db 1664 TTTTATACAAAAATA 1681

RESULT 10
 US-09-764-877-3292
 ; Sequence 3292, Application US/09764877
 ; Patent No. US20020147140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC005
 ; CURRENT APPLICATION NUMBER: US/09/764.877
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 4031
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 3292
 ; LENGTH: 4704
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-877-3292

Query Match 7.2%; Score 18; DB 10; Length 4704;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 74 ATTCCTGCACATACAT 91
 |||||
 Db 1961 ATTCCTGCACATACAT 1978

RESULT 11
 US-09-818-512-3
 ; Sequence 3, Application US/09818512
 ; Patent No. US20020142416A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BEASLEY, Ellen et al.
 ; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001192
 ; CURRENT APPLICATION NUMBER: US/09/818.512
 ; CURRENT FILING DATE: 2001-03-28
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 3
 ; LENGTH: 116592
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(116592)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-818-512-3

Query Match 7.2%; Score 18; DB 10; Length 116592;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 137 TTGTATTAGGTGTATAA 154
 |||||
 Db 94442 TTGTATTAGGTGTATAA 94459

RESULT 12
 US-09-790-988-1/c
 ; Sequence 1, Application US/09790988
 ; Patent No. US20020127687A1
 ; GENERAL INFORMATION:
 ; APPLICANT: SHIGENOBU, SHUJI
 ; APPLICANT: WATANABE, HIDEKI
 ; APPLICANT: HATTORI, MASAHIRA
 ; APPLICANT: SAKAKI, YOSHIYUKI
 ; TITLE OF INVENTION: GENOME DNA OF BACTERIAL SYMBIONT OF APHIDS
 ; FILE REFERENCE: 081356/0159
 ; CURRENT APPLICATION NUMBER: US/09/790.988
 ; CURRENT FILING DATE: 2001-02-23
 ; PRIOR APPLICATION NUMBER: JP2000-107160
 ; PRIOR FILING DATE: 2000-04-07
 ; NUMBER OF SEQ ID NOS: 7
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 640681
 ; TYPE: DNA
 ; ORGANISM: Buchnera sp.
 US-09-790-988-1

Query Match 7.2%; Score 18; DB 10; Length 640681;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 97 TTTTATACAAAAATA 114
 |||||
 Db 483308 TTTTATACAAAAATA 483291

RESULT 13
 US-09-962-436-517/c
 ; Sequence 517, Application US/09962436
 ; Patent No. US20020081301A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Soppet, Daniel
 ; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign

; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-75
; CURRENT APPLICATION NUMBER: US/09/962,436
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US/60/235,082
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/234,924
; PRIOR FILING DATE: 2000-09-25
; NUMBER OF SEQ ID NOS: 568
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 517
; LENGTH: 148
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-962-436-517

Query Match 6.8%; Score 17; DB 10; Length 148;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 109 AAAATAATACACACAA 125
|||||
Db 138 AAAATAATACACACAA 122

RESULT 14
US-09-969-373-1246/c
; Sequence 1246, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Effertz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 1246
; LENGTH: 309
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-1246

Query Match 6.8%; Score 17; DB 10; Length 309;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 149 TTATAATAATAATAAT 165
|||||
Db 267 TTATAATAATAATAAT 251

RESULT 15
US-09-974-300-5415/c
; Sequence 5415, Application US/09974300
; Patent No. US20020146721A1
; GENERAL INFORMATION:
; APPLICANT: Berka, Randy M.
; APPLICANT: Clausen, Ib Groth
; TITLE OF INVENTION: Methods For Monitoring Multiple Gene
; FILE REFERENCE: 10085.500-US
; CURRENT APPLICATION NUMBER: US/09/974,300
; CURRENT FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 09/680,598
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/279,526

; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 8481
; SOFTWARE: FastSeq for Windows version 4.0
; SEQ ID NO 5415
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Bacillus clausii
US-09-974-300-5415

Query Match 6.8%; Score 17; DB 10; Length 433;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 ACATGTTCTTTTAT 104
|||||
Db 32 ACATGTTCTTTTAT 16

Search completed: February 13, 2003, 07:01:29
Job time : 298 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 13, 2003, 03:18:01 ; Search time 2282 Seconds
(without alignments)
1767.167 Million cell updates/sec

Title: US-09-895-435-4
Perfect score: 249
Sequence: 1 999gtaccgcagcattcg.....gacgcgaggagaaatcacatg 249

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estio:*
8: em_hic:*
9: gb_esti:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	22	8.8	796	17 AQ782108	AQ782108 HS_3174_B
2	21	8.4	534	17 CNS00NSF	AL0823117 Arabidops
3	21	8.4	561	17 BH781624	BH781624 fzm021fo
4	21	8.4	584	17 AQ967791	AQ967791 LERIV03TR
5	21	8.4	588	17 AZ527276	AZ527276 266PbE11
6	21	8.4	607	17 AG138789	AG138789 Pan trogl

7	8.4	726	17	BH860734
21	8.4	1180	17	B10174
21	8.0	191	10	AW481697
20	8.0	277	12	BG631395
11	8.0	331	13	B1187613
12	8.0	353	13	B1345004
13	8.0	411	13	B1191005
14	8.0	420	17	AQ058794
15	8.0	462	10	AV682443
16	8.0	718	17	AG174222
17	8.0	818	17	A2538703
18	8.0	920	12	BF275297
19	7.6	263	13	B1744903
20	7.6	382	12	BF396290
21	7.6	386	13	BM406353
22	7.6	399	17	A2741082
23	7.6	401	17	AQ953948
24	7.6	429	14	BM958869
25	7.6	441	9	AU229838
26	7.6	442	17	AQ600207
27	7.6	443	12	BG603695
28	7.6	450	17	AZ175162
29	7.6	453	17	AZ175164
30	7.6	457	17	BH213219
31	7.6	458	13	BM111491
32	7.6	460	17	AZ285657
33	7.6	474	13	BM113619
34	7.6	474	13	BM407225
35	7.6	474	13	BM407275
36	7.6	474	13	BM407323
37	7.6	474	13	BM407330
38	7.6	474	13	BM407331
39	7.6	474	13	BM407401
40	7.6	474	13	BM407411
41	7.6	474	13	BM407418
42	7.6	474	13	BM407443
43	7.6	478	13	BM111174
44	7.6	491	17	CNS060T8
45	7.6	507	13	BM109770

ALIGNMENTS

RESULT 1
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

AQ782108
HS_3174_B1_D08_T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3174 Col=15 Row=H, DNA sequence.
AQ782108
AQ782108.1 GI:5685068
GSS.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 796)
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu

Plate: 3174 row: H column: 15	
Seq primer: T7	
Class: BAC ends	
High quality sequence stop: 796.	
Location/Qualifiers	
1..796	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
/clone="plate=3174 Col=15 Row=H"	
/clone_lib="CIR Approved Human Genomic Sperm Library D"	
/sex="male"	
/note="organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"	
244 a 159 c 169 g 220 t 4 others	
BASE COUNT	
ORIGIN	
Query Match 8.8%; Score 22; DB 17; Length 796;	
Best Local Similarity 100.0%; Pred. No. 8.6;	
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 95 CTTTATTATACAAAAATAAT 116	
DB 591 CTTTATTATACAAAAATAAT 612	
RESULT 2	
CNS00NSF	
LOCUS	
DEFINITION	
CNS00NSF 534 bp DNA linear GSS 28-JUN-1999	
Arabidopsis thaliana genome survey sequence T7 end of BAC P4K3 of	
IGF library from strain Columbia of Arabidopsis thaliana, genomic	
survey sequence.	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
Eukaryota; viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
Rosidae; eurosids II; Brassicales; Arabidopsis.	
1 (bases 1 to 534)	
Salanoubat,M., Choinsne,N., Artiguenave,F., Brottier,P., Wincker,P.,	
Samson,D., Saurin,W., Weissensbach,J. and Quetier,F.	
Unpublished	
2 (bases 1 to 534)	
Genoscope.	
Direct Submission	
AUTHORS	
TITLE	
JOURNAL	
REFERENCE	
Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage :	
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
- Web : www.genoscope.cns.fr)	
Location/Qualifiers	
1..534	
/organism="Arabidopsis thaliana"	
/strain="Columbia"	
/db_xref="taxon:3702"	
/clone="P4K3"	
/clone_lib="IGF"	
/note="end : T7"	
175 a 91 c 85 g 183 t	
BASE COUNT	
ORIGIN	
Query Match 8.4%; Score 21; DB 17; Length 534;	
Best Local Similarity 100.0%; Pred. No. 29;	
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 96 TTTTATTATACAAAAATAAT 116	
DB 15 TTTTATTATACAAAAATAAT 35	
RESULT 3	
BH781624	
LOCUS	
DEFINITION	
BH781624 561 bp DNA linear GSS 28-MAR-2002	
Arabidopsis thaliana genomic clone	
Accession	
Version	
Keywords	
Source	
Organism	
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC	
clade; Panicoidae; Andropogoneae; Zea.	
1 (bases 1 to 561)	
Budiman,M.A., Freese,R.G., Bedell,J.A., Nunberg,A.N. and Lakey,N.D.	
GeneThresher methylation filtered genomic sequences from maize	
unpublished (2002)	
Contact: Bedell JA	
Orion Genomics, LLC	
4041 Forest Park Ave, St. Louis, MO 63108, USA	
Tel: 314 615 6979	
Fax: 314 615 5975	
Email: jbedell@oriongenomics.com	
Plate: fzm021f007 row: c column: 08	
Seq primer: M13 forward	
Class: shotgun	
High quality sequence stop: 561.	
Location/Qualifiers	
1..561	
/organism="Zea mays"	
/cultivar="M017"	
/db_xref="taxon:4577"	
/clone="fzm021f007c08"	
/note="organ: leaf; Vector: pBCSK(-); Site1: HincII; DNA	
end-repaired, size fractionated to enrich for the 0.5 to	
5 kb fraction, ligated into HincII-digested pBCSK(-)	
vector and electroporated into E. coli cells."	
135 a 115 c 147 g 164 t	
BASE COUNT	
ORIGIN	
Query Match 8.4%; Score 21; DB 17; Length 561;	
Best Local Similarity 100.0%; Pred. No. 28;	
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 84 ACATACATGTCCTTTTAT 104	
DB 469 ACATACATGTCCTTTTAT 489	
RESULT 4	
AQ967791	
LOCUS	
DEFINITION	
AQ967791 584 bp DNA linear GSS 28-JAN-2000	
LERIV03TR LERG Arabidopsis thaliana genomic clone LERIV03, DNA	
sequence.	
ACCESSION	
VERSION	
KEYWORDS	
SOURCE	
ORGANISM	
Arabidopsis thaliana	
Eukaryota; viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
Rosidae; eurosids II; Brassicales; Arabidopsis.	
1 (bases 1 to 584)	
Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Utterbach,T.,	
Feldblyum,T., Liang,F., Creasy,T. and Fraser,C.M.	
Genomic survey sequencing of Landsberg erecta ecotype of	
Arabidopsis thaliana and identification of sequence-based	
polymorphisms	
unpublished (2000)	
Contact: Xiaoying Lin	
The Institute for Genomic Research	
9712 Medical Center Dr., Rockville, MD 20850, USA	
Tel: 301 838 0200	
Fax: 301 838 0208	
Email: att@tigr.org	

For additional information, see <http://www.tigr.org/tdb/at/at.html>
 Seq primer: TR
 Class: shotgun.

FEATURES
 source
 1..584
 /organism="Arabidopsis thaliana"
 /strain="Landsberg erecta"
 /db_xref="taxon:3702"
 /clone="LERIV03"
 /clone_lib="LEPG"
 /note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was sheared to 0.4-0.7 kbp before ligation."

BASE COUNT 199 a 99 c 118 g 168 t
 ORIGIN

Query Match 8.4%; Score 21; DB 17; Length 584;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 49 TTAATAATCAGAAATATAT 69
 |||||
 Db 78 TTAATAATCAGAAATATAT 98

RESULT 5
 AZ527276/c
 LOCUS 588 bp DNA linear GSS 07-MAY-2001
 DEFINITION 266PbE11 Pb MBN #21 Plasmodium berghei genomic 3', DNA sequence.
 ACCESSION AZ527276
 VERSION AZ527276.1 GI:13969913
 KEYWORDS GSS.

ORGANISM Plasmodium berghei.
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 588)
 AUTHORS Carlton, J.M.-R. and Dame, J.B.
 TITLE The Plasmodium vivax and P. berghei gene sequence tag projects
 JOURNAL Parasitol. Today (Regul. Ed.) 16 (10), 409 (2000)
 COMMENT Contact: Dame JB
 Dept. of Pathobiology, College of Veterinary Medicine
 University of Florida
 2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
 Tel: 352 392 4700
 Fax: 352 392 9704
 Email: damej@mail.vetmed.ufl.edu
 Seq primer: M13(-20) forward
 Class: shotgun.

FEATURES
 source
 1..588
 /organism="Plasmodium berghei"
 /strain="ANKA clone 15cyl (clone of the ANKA 8417 clone)"
 /db_xref="taxon:5821"
 /clone_lib="Pb MBN #21"
 /dev_stage="asexual blood forms"
 /lab_host="Mus musculus"
 /note="Vector: pBluescript SK(+) vector DNA, phagemid excised from lambda ZAP; Site_1: EcorV; Site_2: EcorV;
 Genomic DNA was prepared from asynchronous blood stage forms of the cloned ANKA isolate of P. berghei grown in laboratory Swiss white mice. The DNA was purified from contaminating host DNA by Hoechst Dye 33258-CsCl ultracentrifugation and precipitated. Purified DNA was digested with mung bean nuclease in the presence of 36-38% formamide at 50 C, as described (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988. Nucleic Acids Research 16:6883-6896). The ends of the digestion fragments were polished using T4 DNA polymerase, and the fragments size selected in the range 500-2000 bp. These were ligated into the EcorV-cleaved and dephosphorylated pBluescript SK(+) vector. Recombinant plasmids were used to transform E. coli XL10-Gold host cells."

BASE COUNT 218 a 74 c 87 g 209 t
 ORIGIN

Query Match 8.4%; Score 21; DB 17; Length 588;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 109 AAAATAATACACACATATT 129
 |||||
 Db 311 AAAATAATACACACATATT 291

RESULT 6
 AG138789/c
 LOCUS 607 bp DNA linear GSS 04-NOV-2001

DEFINITION Pan troglodytes DNA, clone: PTB-153B07.F, genomic survey sequence.
 ACCESSION AG138789
 VERSION AG138789.1 GI:16668467
 KEYWORDS GSS.
 SOURCE Pan troglodytes male lymphoblast DNA, clone_lib:PTB Chimpanzee Male BAC Library clone:PTB-153B07.F.

ORGANISM

Pan troglodytes
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 1
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE BAC end sequences of Library PTB
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 607)
 AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
 TITLE Direct Submission
 JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimpbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/, Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 COMMENT Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
 PRIMERS

Sequencing: -21M13

LIBRARY Vector : pKS145
 R.Site 1 : SacI
 R.Site 2 : SacI
 Location/Qualifiers
 1..607
 /organism="Pan troglodytes"
 /db_xref="taxon:9598"
 /clone="PTB-153B07.F"
 /sex="male"
 /cell_type="lymphoblast"
 /clone_lib="PTB Chimpanzee Male BAC Library"

BASE COUNT 216 a 41 c 142 g 187 t
 ORIGIN

Query Match 8.4%; Score 21; DB 17; Length 607;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 96 TTTTATATACAAAAATAAT 116
 |||||
 Db 333 TTTTATATACAAAAATAAT 313

RESULT 7
 BH860734
 LOCUS 726 bp DNA linear GSS 08-JUL-2002

DEFINITION ORNL233 Poplar BAC Library Populus balsamifera subsp. trichocarpa genomic, DNA sequence.
 ACCESSION BH860734
 VERSION BH860734.1 GI:21711555
 KEYWORDS GSS.

```

SOURCE
ORGANISM      Populus balsamifera subsp. trichocarpa.
               Populus balsamifera subsp. trichocarpa
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eurosids I; Malpighiales; Salicaceae; Populus.

REFERENCE
AUTHORS       Stirling,B., Yang,Z., Gunter,L., Tuskan,G. and Bradshaw,H.D.
TITLE         Comparative sequence analysis between orthologous regions of the
               Arabidopsis and Populus (poplar) genomes reveals substantial
               synteny and microcollinearity
               Unpublished (2002)
JOURNAL
COMMENT       Contact: Stirling B
               Department of Molecular Biology
               Immunex Corporation
               51 University Street, Seattle, WA 98101 USA
               Tel: 206 587 0430 ext. 34429
               Email: StirlingB@immunex.com
               Insert Length: 726 Std Error: 0.00
               Class: shotgun.

FEATURES
source
Location/Qualifiers
1..726
/organism="Populus balsamifera subsp. trichocarpa"
/cultivar="Nisqually-1"
/db_xref="taxon:3694"
/clone_lib="Poplar BAC Library"
/sex="female"
/tissue_type="leaf"
/lab_host="E.coli DH10B"
/note="Vector: pBeloBACII; Site_1: HindIII; Site_2:
HindIII; Ramets of P. trichocarpa (clone 383-2499) were
grown in the greenhouse. The trees were placed in the
dark for two days to deplete starch reserves prior to
harvesting approximately 60 g of fresh, young leaves.
Extraction of high molecular weight genomic DNA from leaf
nuclei was performed as described in Zhang et al. (1994).
Poplar genomic DNA was partially digested with HindIII
followed by three rounds of size selection from agarose
gels following separation by pulsed field gel
electrophoresis (PFGE). In the first two size selections,
the pulsed field gel was run with a 90 second pulse at
160 V for a total run time of 18 hours at 110C. In both
runs DNA ranging from 200 kb to 400 kb was excised from
the gel and used in the next size selection. The final
size selection gel was run using a 6 second pulse at 150V
for 11 hours at 110C, and the compressed band
representing DNA fragments greater than 150 kb was
excised. Prior to ligation the final size-selected DNA
was released from agarose by electroelution. The agarose
gel slice containing the >150 kb DNA was fragmented with
a razor blade and the resulting pieces placed in dialysis
bags (Gibco BRL, Rockville, MD). Electroelution was
carried out at 110C for 2 hours at 200 V followed by a 90
second reversed-current pulse. Eluted DNA was quantified
by inspection on an agarose gel. HindIII-digested and
dephosphorylated pBeloBAC II (Shizuya et al. 1992) was
used for library construction. A molar ratio of
approximately 3:1 vector:insert was used for ligation.
Transformations were performed by electroporation using
Gene Hogs electrocompetent cells (Research Genetics,
Huntsville, AL). Transformed cells were plated onto LB
(Difco, Sparks, MD) agar supplemented with 12.5 mg/ml
chloramphenicol and 0.25% X-gal. White colonies (N=3D 48
,384) were picked and grown in 126 384-well microtiter
plates."
BASE COUNT    227 a 101 c 118 g 280 t
ORIGIN
Query Match      8.4%; Score 21; DB 17; Length 726;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

94 TCTTTTATACAAAAATA 114
|||||

SOURCE
ORGANISM      B10174
               F15G16-Sp6 IGF Arabidopsis thaliana genomic clone F15G16, DNA
               sequence.
               1180 bp
               DNA linear GSS 14-MAY-1997
DEFINITION
LOCUS          B10174
ACCESSION     B10174
VERSION       B10174.1 GI:2091293
KEYWORDS      GSS.
SOURCE        thale cress.
ORGANISM      Arabidopsis thaliana
               Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
               Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
               Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS       Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
               Ecker,J.
TITLE         BAC End Sequences at ATGC
JOURNAL
COMMENT       Unpublished (1997)
               Other_GSSs: F15G16-T7.1
               Contact: Ecker J.
               Arabidopsis Thaliana Genome Center
               University of Pennsylvania
               Dept. of Biology, University of Pennsylvania, Philadelphia, PA
               19104
               Tel: 215-898-9384
               Fax: 215-898-8780
               Email: jecker@genome.bio.upenn.edu
               Seq primer: Sp6
               Class: BAC ends
               High quality sequence start: 70
               High quality sequence stop: 646.

FEATURES
source
Location/Qualifiers
1..1180
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone_lib="F15G16"
/clone_lib="IGF"
/sex="hermaphrodite"
/note="Vector: pBeloBACII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altman"
BASE COUNT    337 a 267 c 197 g 373 t
ORIGIN
Query Match      8.4%; Score 21; DB 17; Length 1180;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

96 TTTTATACAAAAATAAT 116
|||||

Db 132 TTTTATACAAAAATAAT 152

SOURCE
ORGANISM      AW481697
               37330 MARC 1BOV Bos taurus cDNA 5', mRNA linear EST 25-APR-2001
LOCUS          AW481697
DEFINITION     37330 MARC 1BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION     AW481697
VERSION       AW481697.1 GI:7051803
KEYWORDS      EST.
SOURCE        cow.
ORGANISM      Bos taurus
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
               Bovidae; Bovinae; Bos.
REFERENCE
AUTHORS       Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
               Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
               G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
               Perlea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and

```

TITLE
Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL
Genome Res. 11 (4), 626-630 (2001)
MEDLINE
21180013
COMMENT
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers
FORWARD: AGGAACACGCTATGACCAT
BACKWARD: GTTTCACGATCAGCAGC
Plate: 22 row: D column: 4
Seq primer: ATTTAGGTGACACTATAG.

FEATURES
Location/Qualifiers

1..191
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC IBOV"
/tissue_type="pooled"
/lab_host="DH108"
/note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from lymph node, ovary,
fat, hypothalamus, and pituitary."

BASE COUNT 56 a 42 c 30 g 63 t
ORIGIN
Query Match 8.0%; Score 20; DB 10; Length 191;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 TTTTATACAAAAATAA 115
|||||
Db 182 TTTTATACAAAAATAA 163

RESULT 10 277 bp mRNA linear EST 19-APR-2001
BG631395
LOCUS
DEFINITION
cC-esf1cLEL7K03a1 Tomato flower library from a mixture of
developmental stages Lycopersicon esculentum cDNA clone

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 277)
van der Hoeven,R.S. and Tanksley,S.D.
ESTs from a tomato flower library
Unpublished (2001)
Contact: Rutgers S. van der Hoeven
Cornell University
252 Emerson Hall, Ithaca, NY 14850, USA
Tel: 607 255 7886
Fax: 607 255 6683
Email: rv19@cornell.edu
3 prime sequence.

LOCATION/Qualifiers

1..277
/organism="Lycopersicon esculentum"
/cultivar="E6203"
/db_xref="taxon:4081"
/clone="cC-esf1cLEL7K03a1"

FEATURES
Source

/clone_lib="Tomato flower library from a mixture of
developmental stages"
/tissue_type="developing flower buds and open flowers"
/dev_stage="4-8 week old plants"
/lab_host="XL0LR"
/note="Vector: pBK.CMV; Site_1: EcoRI; Site_2: XhoI;
Flowers and flower buds were collected from greenhouse
grown plants and used for library construction (cLEL)."
BASE COUNT 92 a 47 c 34 g 104 t
ORIGIN

Query Match 8.0%; Score 20; DB 12; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 TTTATACAAAAATAATACA 119

|||||

Db 39 TTTATACAAAAATAATACA 58

RESULT 11

BI187613

LOCUS

DEFINITION

BI187613 331 bp mRNA linear EST 10-JUL-2001
bib01fs.f1 Fusarium sporotrichioides Tri 10 overexpressed cDNA
library Fusarium sporotrichioides cDNA clone bib01fs 3', mRNA
sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Fusarium sporotrichioides.
Fusarium sporotrichioides
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreales; mitosporic Hypocreales; Fusarium.

REFERENCE

AUTHORS

Ren,Q., Tag,A., Pielow,A., Lai,H., Kupfer,C., Peterson,A., Beremand
,M. and Roe,B.

TITLE

JOURNAL

COMMENT

Analysis of a Fusarium sporotrichioides EST database
Unpublished (2001)
Other_ESTs: bib01fs.r1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu

Contact Dr. Marian Beremand regarding clone availability Included
is the best homolog from a blastx search of Genbank nr 04-09-01
68 5.5 gi|13376187|ref|NP_0 hypothetical protein FLJ23554
[Homosapiens

Seq primer: M13-20

High quality sequence stop: 318.

FEATURES

Location/Qualifiers

Source

1..331

/organism="Fusarium sporotrichioides"

/strain="Tri 10"

/db_xref="taxon:5514"

/clone="bib01fs"

/clone_lib="Fusarium sporotrichioides Tri 10 overexpressed

cDNA library"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:

XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript

; 3' end of cDNA cloned into XhoI site of pBluescript"

BASE COUNT 92 a 77 c 69 g 93 t
ORIGIN

Query Match 8.0%; Score 20; DB 13; Length 331;

Best Local Similarity 100.0%; Pred. No. 1e+02;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 AGAGTTAAAAAATCAGAAAA 64

|||||

Db 309 AGAGTTAAAAAATCAGAAAA 328

JOURNAL
COMMENT

Unpublished (2001)
Other_ESTs: j2f05fs.r1
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Contact Dr. Marian Beremand regarding clone availability Included
is the best homolog from a blastx search of Genbank nr 04-09-01
71 3.6 gll13376187|ref|NP_0 hypothetical protein FLJ23554
[Homosapiens
Seq primer: M13-20
High quality sequence stop: 313.
Location/Qualifiers
1..411
/organism="Fusarium sporotrichioides"
/strain="Tri 10"
/db_xref="taxon:5514"
/clone="j2f05fs"
/clone_lib="Fusarium sporotrichioides Tri 10 overexpressed
cDNA library"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; 5' end of cDNA cloned into EcoRI site of pBluescript
; 3' end of cDNA cloned into XhoI site of pBluescript"
BASE COUNT 115 a 99 c 85 g 112 t
ORIGIN

FEATURES
Source

Query Match 8.0%; Score 20; DB 13; Length 411;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 AGAGTTAAAAATCAGAAA 64
|||||
Db 301 AGAGTTAAAAATCAGAAA 320
|||||
RESULT 14
LOCUS AQ058794/c 420 bp DNA linear GSS 20-APR-1999
DEFINITION RPC111-52A23.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-52A23,
DNA sequence.
ACCESSION AQ058794
VERSION AQ058794.1 GI:3348755
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 420)
AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and
Venter,J.C.
TITLE Use of BAC End Sequences for Sequence-Ready Map Building (1998)
JOURNAL Unpublished (1998)
COMMENT Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html
Class: BAC ends.
Location/Qualifiers
1..420
/organism="Homo sapiens"

RESULT 12
BI345004

LOCUS BI345004 353 bp mRNA linear EST 30-JUL-2001
DEFINITION 373681 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI345004
VERSION BI345004.1 GI:15038284
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 353)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E.,
Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W.
and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for
EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCGCCAGTCAGCAG
Plate: 121 row: I column: 6
Seq primer: ATTTAGGTGACACTATAG.
Location/Qualifiers
1..353
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2P1G"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: PCMV SPORT6; Site_1: NotI; Site_2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 80 a 80 c 98 g 95 t
ORIGIN

FEATURES
Source

Query Match 8.0%; Score 20; DB 13; Length 353;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 93 TTCTTTTATACAAAAA 112
|||||
Db 259 TTCTTTTATACAAAAA 278
|||||
RESULT 13
BI191005
LOCUS BI191005 411 bp mRNA linear EST 10-JUL-2001
DEFINITION BI191005.f1 Fusarium sporotrichioides Tri 10 overexpressed cDNA
library Fusarium sporotrichioides cDNA clone j2f05fs 3', mRNA
sequence.
ACCESSION BI191005
VERSION BI191005.1 GI:14664684
KEYWORDS EST.
SOURCE Fusarium sporotrichioides.
ORGANISM Fusarium sporotrichioides.
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Hypocreales; mitosporic Hypocreales; Fusarium.
REFERENCE 1 (bases 1 to 411)
AUTHORS Ren,Q., Tag,A., Peplow,A., Lai,H., Kupfer,C., Peterson,A., Beremand
M. and Roe,B.
TITLE Analysis of a Fusarium sporotrichioides EST database

Job time : 2292 secs

```

/db_xref="GDB:7519606"
/db_xref="taxon:9606"
/clone="RPCI-11-52A23"
/clone_lib="RPCI-11"
/sex="Male"
/cell_type="Lymphocytes"
/notes="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
RPCI11 Human Male BAC Library"
BASE COUNT      122 a      84 c      63 g      151 t
ORIGIN

```

```

Query Match      8.0%; Score 20; DB 17; Length 420;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 146 GTGTTATATAATAATAAT 165
      |||||||||||||||||||
DB 340 GTGTTATATAATAATAAT 321

```

```

RESULT 15
AV682443
LOCUS          AV682443      462 bp      mRNA      linear      EST 16-JAN-2002
DEFINITION    AV682443 GKB Homo sapiens cDNA clone GKBAD411 5', mRNA sequence.
ACCESSION     AV682443
VERSION       AV682443.1 GI:10284306
KEYWORDS      EST.
SOURCE        human.
ORGANISM      Homo sapiens

```

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

XU,X., HUANG,J., XU,Z., QIAN,B., ZHU,Z., YAN,Q., CAI,T., ZHANG,X.,
XIAO,H., QU,J., LIU,F., HUANG,Q., CHENG,Z., LI,N., DU,J., HU,W.,
SHEN,K., LU,G., FU,G., ZHONG,M., XU,S., GU,W., HUANG,W., ZHAO,X.,
HU,G., GU,J., CHEN,Z., and HAN,Z.

```

```

Insight into hepatocellular carcinogenesis at transcriptome level
by comparing gene expression profiles of hepatocellular carcinoma
with those of corresponding noncancerous liver
Proc. Natl. Acad. Sci. U.S.A. 98 (26), 15089-15094 (2001)
21625106

```

```

Contact: Zequang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

```

This clone is available at CHGC in Shanghai.

```

FEATURES
    source
        1..462
            Location/Qualifiers
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="GKBAD411"
                /clone_lib="GKB"
                /tissue_type="hepatocellular carcinoma"
                /dev_stage="Adult"
                /lab_host="SOLR"
                /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
                XhoI"
BASE COUNT      186 a      70 c      58 g      148 t
ORIGIN

```

```

Query Match      8.0%; Score 20; DB 10; Length 462;
Best Local Similarity 100.0%; Pred. No. 88;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 100 TTTATACAAAAATAATACA 119
      |||||||||||||||||||
DB 368 TTTATACAAAAATAATACA 387

```

Search completed: February 13, 2003, 06:30:44

THIS PAGE BLANK (USPTO)